

Luciana Lucente

Aspectos Dinâmicos da Fala e da Entoação no Português Brasileiro

Dynamic Aspects of Speech and Intonation in Brazilian Portuguese



Campinas

2012

Universidade Estadual de Campinas
Instituto de Estudos da Linguagem - IEL

Aspectos Dinâmicos da Fala e da Entoação no Português Brasileiro

Luciana Lucente

Orientador/Supervisor Dr. Plínio Almeida Barbosa

Dynamic Aspects of Speech and Intonation in Brazilian Portuguese

Tese de Doutorado apresentada no
Instituto de Estudos da Linguagem
como parte dos requisitos para obtenção
do título de Doutora em Linguística.

*A Thesis presented at the Institute of
Language Studies in partial fulfillment
of the requirements for the degree of
Ph.D. in the subject of Linguistics.*

Campinas, 2012

FICHA CATALOGRÁFICA ELABORADA POR
CRISLLENE QUEIROZ CUSTODIO – CRB8/8624 - BIBLIOTECA DO INSTITUTO DE
ESTUDOS DA LINGUAGEM – UNICAMP

L963a

Lucente, Luciana, 1978-

Aspectos dinâmicos da fala e da entoação do português brasileiro / Luciana Lucente. -- Campinas, SP : [s.n.], 2012.

Orientador : Plínio Almeida Barbosa.

Tese (doutorado) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem.

1. Sistemas dinâmicos. 2. Entoação (Fonética). 3. Notação entoacional. 4. Fala espontânea (Linguística). 5. Análise do discurso. I. Barbosa, Plínio Almeida, 1966-. II. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.

Informações para Biblioteca Digital

Título em inglês: Dynamic aspects of speech and intonation in brazilian portuguese.

Palavras-chave em inglês:

Dynamical systems

Intonation (Phonetics)

Intonational annotation

Spontaneous speech (Linguistics)

Discourse analysis

Área de concentração: Linguística.

Titulação: Doutora em Linguística.

Banca examinadora:

Plínio Almeida Barbosa [Orientador]

Maria Bernadete Marques Abaurre

Sandra Madureira Fontes

Pablo Arantes

Miguel José Alves de Oliveira Junior

Data da defesa: 13-08-2012.

Programa de Pós-Graduação: Linguística.

A Comissão Julgadora dos trabalhos de Defesa de Tese de Doutorado , em sessão pública realizada em 13 de agosto de 2012, considerou o candidato LUCIANA LUCENTE aprovado.

BANCA EXAMINADORA:

Plínio Almeida Barbosa

Maria Bernadete Marques Abaurre

Sandra Madureira Fontes

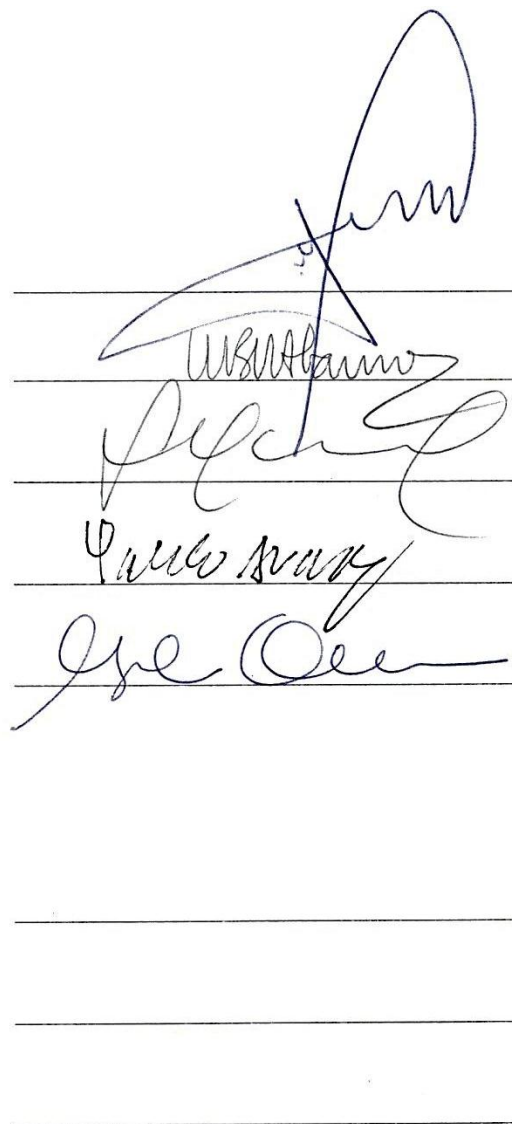
Pablo Arantes

Miguel José ALves de Oliveira Junior

Maria Filomena Spatti Sândalo

Tommaso Raso

João Antonio Moraes



The image shows five handwritten signatures in blue ink, each written over a horizontal line. The signatures are: 1. Plínio Almeida Barbosa (top), 2. Maria Bernadete Marques Abaurre, 3. Sandra Madureira Fontes, 4. Pablo Arantes, and 5. Miguel José Alves de Oliveira Junior (bottom). The signatures are written in a cursive style.

IEL/UNICAMP
2012

à minha avó, Maria Pereira, que me ensinou a ler e a escrever

AGRADECIMENTOS

Meus agradecimentos às pessoas e instituições que me deram suporte no percurso desta tese se dividem entre academia, amigos e família.

Inicio esses agradecimentos pela Fapesp (Fundação de Amparo à Pesquisa do Estado de São Paulo) pelo suporte financeiro, sem o qual a realização da minha pesquisa não seria possível. Agradeço à Unicamp, universidade em que passei mais de 10 anos de minha vida, estudando, pesquisando e fazendo amigos. Agradeço à Unicamp pela formação de extrema qualidade que recebi no IEL (Instituto de Estudos da Linguagem), pelos atendimentos médico (Cecom) e psicológico (Sappe), pela bolsa trabalho (SAE) que tive logo que ingressei na universidade, enfim, pelo suporte que a Unicamp pode proporcionar ao aluno e que deveria ser estendido a todas as universidades públicas, e não o contrário.

Ainda na vida acadêmica agradeço à convivência com os amigos do Grupo de Estudos de Prosódia da Fala, Pablo Arantes, Aveliny Lima-Gregio, Sandra Merlo e Ana Carolina Constantini, pessoas muito queridas e com quem pude trocar experiências de vida e de trabalho. Agradeço à profa. Sandra Madureira pela presença constante nos nossos encontros, aos professores do IEL, Wilmar D'Angelis e Edson Françoso, pela influência na minha forma de trabalhar, e à professora Bernadete Abaurre, pelas observações feitas no exame de qualificação. Agradeço ainda à professora Julia Hirschberg, que me recebeu muito abertamente (para não dizer maternalmente) na Universidade de Columbia, e em sua própria casa, durante o estágio que realizei nos Estados Unidos.

E agradeço ao meu orientador, Plínio Barbosa, que está para mim entre academia, amizade e família, que me acompanhou-orientou-aconselhou durante oito anos de uma forma que não consigo pensar que poderia ter sido melhor e que me mostrou confiança ao me delegar tarefas de bastante responsabilidade. Uma pessoa sensível, honesta, justa, divertida, amante da vida, da natureza, da poesia, da ciência. Um verdadeiro mestre.

Agradeço aos amigos diretamente ligados a essa tese em alguma momento, seja auxiliando no trabalho ou apenas sendo amigos, são eles: Joseane Porfírio, André Xavier,

Cristiane Silva, Luana Campos, Lílían Teixeira, Lívía Cucatto, Thalita Valle, Jael Gonçalves e Meghan Armstrong, amiga de aventuras linguísticas pelo mundo afora. Amigos que conheci na Unicamp, ou por meio do doutorado, e os que vêm da infância: Leandro Dinis, Marianne Pesci de Matos, Aline Daltro, Rodrigo Guimarães, Ana Luiza Lopes, Carolina Assis, Mahayana Godoy, Marcelo Bosso, Fernanda Consoni, Marcus Avelar, Lisa Barri, Joana Maria, José Cláudio Moreno, Laura Batista e Simone Palacini.

Os amigos-sujeitos, que sempre estavam dispostos a participar dos meus experimentos: Aline Machado, Aroldo Andrade, Bebel Debien, Fernando Gregio, Jordana Reis, Ricardo Souza, Juliana Gobbo, Leandro Vendemiatti, Letícia Rebollo Couto, Lílían Costa, Luciana Falasca, Marcelo Porqueres, Marcos Pires, Pablo Faria e Wellington Silva. E um agradecimento especial à Ana Carolina Constantini e à Joseane Porfírio que, de tanto ajudar, com o passar do tempo já tinham certa intimidade com o corpus!

Agradeço aos meus pais, Francisco e Maria Elena Lucente, por terem me educado de uma forma livre. Livre para que eu pudesse escolher meus caminhos de forma independente e adulta, e por terem me apoiado nas minhas decisões sem nunca tentar me influenciar ou dissuadir. Minha irmã Gisela, que esteve ao meu lado desde a escolha pelo curso da graduação, passando pela matrícula na Unicamp até ir me encontrar emocionada no *Columbus Circle*. Agradeço à Érica Andrade Figueirêdo, que esteve comigo nesses tantos anos de Unicamp, me apoiando, me encorajando, me animando em momentos de desânimo, me fazendo rir, fazendo minha vida mais feliz.

E por fim, agradeço ao amado e fiel companheiro felino João Batista, linguista por indução, sempre presente entre os livros e o computador. E ao fiel Sampson, meu “melhor amigo do homem” em Nova York.

*Two roads diverged in a Wood, and I –
I took the one less traveled by,
And that has made all the difference.*

(Robert Frost, *The Road Not Taken*)

Resumo

Esta tese explora a relação entre padrões entoacionais ritmo e discurso de acordo com o programa de investigação dos sistemas dinâmicos. O estudo dessas relações foram feitas tendo como base o Modelo Dinâmico do Ritmo da Fala, proposto por Barbosa (2006), o Sistema DaTo de notação entoacional, proposto por Lucente (2008) e o Modelo Computacional da Estrutura do Discurso, proposto por Grosz & Sidner (1986). O Modelo de Dinâmico do Ritmo sugere que o ritmo da fala seja resultado da ação de dois osciladores – um acentual e outro silábico – que ao receberem na entrada do sistema informações de níveis linguísticos superiores e de uma pauta gestual, geram a duração gestual na saída. A hipótese desta tese é que, paralelamente a esses osciladores, um oscilador glotal possa agir controlando os padrões entoacionais da fala. Os padrões, ou ciclos entoacionais, em que se organiza a entoação do PB emergem quando relacionados à segmentação de trechos de discurso em modalidade espontânea. Para cada trecho de fala classificado como espontâneo de acordo com um critério proposto nesta tese, o discurso é segmentado no sistema DaTo em unidades linguisticamente estruturadas, que contém os propósitos de comunicar e atrair atenção. Cada um destes segmentos do discurso se alinham a um padrão entoacional iniciado por um contorno entoacional ascendente (LH ou >LH) e finalizado por um contorno descendente (LHL) ou por um nível de fronteira baixo (L). Alinhado a este padrão formado entre entoação e discurso está também o ritmo. Com o acréscimo de uma camada no sistema DaTo para a segmentação dos enunciados em grupos acentuais pôde-se observar o alinhamento entre a segmentação dos grupos acentuais e a notação dos contornos entoacionais coincidindo com fronteiras das unidades discursivas. A observação do alinhamento entre entoação, ritmo e discurso tendo como atratores os grupos acentuais possibilitou a proposta de inserção de um oscilador glotal ao Modelo Dinâmico do Ritmo.

Palavras-chave: sistemas dinâmicos; entoação; notação entoacional; ritmo; fala espontânea; análise do discurso.

Abstract

This thesis explores the relationship between intonational patterns and its relationship with speech rhythm and discourse, according to the dynamic systems research program. The study of these relationships were based on Barbosa's (2006) Dynamic Model of Speech Rhythm; on DaTo intonational annotation system proposed by Lucente (2008); and on the Computational Model of the Structure of Discourse, proposed by Grosz & Sidner (1986). The Dynamic Model of Rhythm suggests that speech rhythm is the result of two oscillators action – accentual and syllabic - which receive linguistic and gestural information as input, and give the gestural duration as output. This thesis hypothesis is that in addition to these oscillators, a glottal oscillator can act controlling the intonation patterns of speech. These patterns, or intonational cycles, which organize the BP intonation, emerge when related to the spontaneous discourse segmentation. For each discourse segment classified as spontaneous, according to a criteria proposed in this thesis, the speech is segmented into the DaTo system framework in linguistically structured units, which contains the purposes of communication and attention. Each of these segments is aligned to the speech intonation pattern delimited by a rising contour (LH or > HL) at the beginning and by a falling contour (LHL), or a boundary level (L), at the end. The speech rhythm is also aligned to the pattern formed between intonation and discourse. By the inclusion of a new layer for the stress groups segmentation into DaTo system was possible to observe the alignment between stress group segmentation and intonational annotation coinciding with discourse segments boundaries. The alignment between intonation, rhythm and discourse, having the stress groups as attractors, allowed us to propose the insertion of a glottal oscillator into the Dynamic Model of Rhythm.

Key-words: dynamic systems; intonation; intonational annotation; rhythm; spontaneous speech; discourse analysis.

SUMÁRIO

Lista de Figuras.....	xxi
Lista de Quadros.....	xxii
Lista de Tabelas	xxii
Simbologia de Transcrição do Corpus	xxii
INTRODUÇÃO.....	1
I ENTOAÇÃO E TEORIA DINÂMICA	3
1. ENTOAÇÃO E TEORIA DINÂMICA	5
1.1 Entoação	5
1.2 O programa dinâmico.....	6
1.2.1 O Programa de Investigação Científica segundo Lakatos	10
1.3 Modelo Dinâmico do Ritmo	14
1.4 Sumário.....	16
II FALA E DISCURSO.....	17
1. CORPUS DE FALA	19
1.1 Em defesa da fala espontânea	21
1.2 Classificação da fala espontânea: gênero e controle.....	25
1.2.1 Classificação do corpus.....	26
1.2.2 Teste de Percepção.....	28
1.3 Sumário.....	32
2. ESTRUTURA DO DISCURSO	33
2.1 Estrutura do Discurso - Modelo de Grosz e Sidner.....	33
2.1.1 Estrutura Linguística	35
2.1.2 Estrutura das Intenções.....	36
2.1.3 Estado de Atenção	39
2.2 Aplicação do Modelo	41
2.3 Elementos de Construção do Discurso.....	43
2.3.1 Interrupções.....	44
2.3.1.1 Hesitação	45

2.3.1.2 Repetição	46
2.3.1.3 Correção	48
2.3.1.4 Parafraseamento	49
2.3.1.5 Parentetização	50
2.3.2 Elementos de construção do discurso sem interrupção	52
2.3.2.1. <i>Status</i> da informação	52
2.3.2.2 Referenciação	55
2.3.2.3 Marcadores Discursivos	56
2.4 Discussão	58
2.5 Sumário	58
 III ASPECTOS DINÂMICOS DA FALA E DA ENTOAÇÃO	61
1. NOTAÇÃO E DESCRIÇÃO ENTOACIONAL	63
1.1 Teoria MA	64
1.2 Perspectiva Dinâmica	68
1.3 Notação DaTo	73
1.3.1 Contornos Dinâmicos	74
1.3.1.1 Teste de Percepção: contornos ascendentes e descendentes	75
1.3.1.1.2 Contornos ascendentes: LH, >LH, vLH e HLH	76
i) <i>LH - rising contour</i>	76
ii) <i>>LH - late rising</i>	77
iii) <i>Experimento de Percepção LH, >LH: alinhamento e altura</i>	78
iv) <i>vLH - compressed rising</i>	83
v) <i>HLH - falling-rising ou “contorno do camelo”</i>	85
1.3.1.2 Contornos descendentes LHL, HL, >HL e vHL	86
i) <i>LHL- rising-falling</i>	87
ii) <i>HL - falling</i>	88
iii) <i>Experimento de percepção HL, LHL: Alinhamento e altura</i>	89
i) <i>> HL - late falling</i>	92
v) <i>> vHL- compressed falling</i>	95
1.3.1.3 Níveis de Fronteira	96
1.3.2 Transcrição Ortográfica	97
1.3.3 Notação Pragmática	98

1.3.4 Unidades V-V	99
1.3.5 Grupos Acentuais	100
1.3.6 Segmentos Discursivos	101
1.4 Sumário	101
2. INTERAÇÃO DISCURSO – RITMO - ENTOÇÃO	103
2.1 Interações entre contorno entoacional, ritmo e discurso	103
2.2 Nova Hipótese para o Modelo Dinâmico do Ritmo	110
2.3 Sumário	113
IV CONCLUSÃO	115
CONCLUSÃO	117
REFERÊNCIAS BIBLIOGRÁFICAS	119
Anexo I	127
Anexo II	145
Anexo III	157
Anexo IV	173
Anexo IV	177

Lista de Figuras

Figura 1	7
Figura 2	7
Figura 3	15
Figura 4	27
Figura 5	29
Figura 6	30
Figura 7	31
Figura 8	37
Figura 9	40
Figura 10	43
Figura 11	65
Figura 12	67
Figura 13	69
Figura 14	70
Figura 15	70
Figura 16	71
Figura 17	72
Figura 18	73
Figura 19	77
Figura 20	78
Figura 21	78
Figura 22	80
Figura 23	81
Figura 24	81
Figura 25	84
Figura 26	85
Figura 27	86
Figura 28	87
Figura 29	88
Figura 30	90
Figura 31	93
Figura 32	94
Figura 33	94
Figura 34	95
Figura 35	96
Figura 36	97
Figura 37	98
Figura 38	100

Figura 39	104
Figura 40	105
Figura 41	106
Figura 42	108
Figura 43	109
Figura 44	111

Lista de Quadros

Quadro 1	36
Quadro 2	38
Quadro 3	38
Quadro 4	41

Lista de Tabelas

Tabela 1	20
Tabela 2	42
Tabela 3	66
Tabela 4	97
Tabela 5	105
Tabela 6	107
Tabela 7	109

Simbologia de Transcrição do Corpus

[]	sobreposição/turno de outro falante
/	interrupção
:	pausa até 20ms
::	pausa até 35ms
:::	pausa até 50ms
::::	pausa até 65ms
:::::	pausa > 65ms

INTRODUÇÃO

Esta tese, que se propõe a estudar os aspectos dinâmicos da fala e da entoação no português brasileiro, é um desdobramento da dissertação de mestrado “DaTo: um sistema de notação entoacional do português brasileiro baseado em princípios dinâmicos. Ênfase no foco e na fala espontânea” (Lucente, 2008), que propôs um sistema de notação entoacional para o português brasileiro (doravante PB).

O desdobramento do estudo anterior, que se limitava a descrever padrões para validar uma notação da entoação do PB, prevê uma análise aprofundada dos aspectos comunicativos e pragmáticos envolvidos na entoação. Para tal investigação foi adotada uma perspectiva dinâmica, ou um programa dinâmico de investigação científica, de acordo com a filosofia da ciência de Lakatos.

O estudo da entoação a partir de uma perspectiva dinâmica assume a interação de aspectos físicos e discursivos e, sobretudo, assume o fato de que a fala é parte de um sistema cognitivo-motor para o qual é impossível fazer uma separação entre um sistema físico e um sistema de intenções. A relação entre a produção e a organização da entoação e do ritmo da fala bem como a relação do estudo da entoação com o discurso vão mais longe que as abordagens anteriores ao mostrar a simetria existente entre os sistemas físico e cognitivo.

Para abordar esses aspectos a tese se divide em duas partes distintas, uma dedicada ao estudo do discurso e outra ao estudo da entoação.

O estudo do discurso segue um modelo computacional que pretende dar conta da estrutura discursiva baseando-se em estruturas menores, como a intenção e a atenção dos participantes de uma conversa. Por meio de elementos construtores do discurso, propõe-se uma segmentação hierárquica dos segmentos discursivos, a partir dos quais se pode mapear a estrutura das intenções.

A segunda parte da tese se concentra na análise entoacional de trechos de discursos espontâneos em PB – segundo uma classificação proposta nesta tese – por meio do sistema DaTo. Os resultados dessa análise entoacional indicam a existência de uma

estrutura dinâmica, que abrange além do discurso e da entoação, também aspectos do ritmo do PB, já abordados nos estudos de Barbosa (2006; 2007; 2008; 2010). O modelo adotado para a análise da sincronia entre discurso, entoação e ritmo é o Modelo Dinâmico do Ritmo da Fala (Barbosa, 2006), para o qual se apresenta uma reformulação redefinindo a interação entre o oscilador glotal e os demais osciladores, silábico e acentual. Isso foi possível após a análise dos dados de texto e áudio do corpus de fala espontânea. Esse percurso se desenvolverá em três capítulos.

O primeiro capítulo, “Entoação e Teoria Dinâmica”, descreve a teoria de sistemas dinâmicos e o modelo teórico adotado nesta tese, procurando discutir a existência de um programa dinâmico de investigação científica.

O segundo capítulo, intitulado “Fala e Discurso”, se divide em duas partes, a primeira descrevendo a composição e as propriedades do corpus de fala utilizado e propondo uma forma de classificação da fala espontânea. A segunda parte discutindo um modelo de análise e segmentação do discurso, utilizado na segmentação do corpus, baseado nos componentes de atenção, intenção e estrutura linguística.

O terceiro capítulo, “Aspectos Dinâmicos da Fala e da Entoação”, também se divide em duas partes, a primeira rediscute o sistema DaTo de notação entoacional e a segunda mostra como o discurso, a entoação e o ritmo interagem de forma dinâmica. Por fim, uma última parte é reservada à conclusão da pesquisa.

IDENTIFICAÇÃO E TEORIA DINÂMICA

1. ENTOAÇÃO E TEORIA DINÂMICA

1.1 Entoação

A entoação comumente aparece relacionada a aspectos físicos do contorno entoacional. Em um sentido estreito ela “*se refere ao uso de características fonéticas suprasegmentais para expressar significados pragmáticos no nível da sentença de forma linguisticamente estruturada*” (Ladd, 1996a, p. 6) e é um dos componentes da prosódia, termo mais amplo que inclui também a descrição de características dinâmicas e temporais associadas aos aspectos formais (Barbosa, 2009).

As características suprasegmentais a que se refere (1996a) são a frequência fundamental (doravante f_0), intensidade e duração, os correspondentes de ordem fonética dos fenômenos psicofísicos, de *pitch*, volume (*loudness*) e duração (*length*) (Lehiste, 1970). Dentre essas características, f_0 é o correlato direto do aspecto fonético que a entoação assume nos estudos prosódicos. A f_0 é medida em Hz (*Hertz*) e corresponde ao número de vezes por segundo em que as pregas vocais completam um ciclo de vibração. Esses ciclos são controlados pelos músculos da laringe que determinam a tensão nas pregas vocais, como também por forças aerodinâmicas do sistema respiratório sublaríngeo (Botinis et al, 2001). Essa definição de entoação que envolve elementos físicos e psicofísicos, ou fonéticos e fonológicos, na sua produção faz com esta possa ser vista como um componente dinâmico da fala.

Na concepção clássica de Saussure (1916), a fala se contrapõe ao conceito de língua. Enquanto a língua é um produto de natureza social não constituindo uma função do falante, a fala se constitui como um ato individual de vontade e inteligência. Como um ato individual, a fala possui caráter sincrônico, refletindo fenômenos e usos em um dado momento, enquanto a língua se volta para a diacronia com seu caráter evolutivo.

Dessa forma, o conceito de entoação se liga à fala e à língua, expressando as funções comunicativas presentes na fala do indivíduo que estabelecem os padrões entoacionais que dizem respeito à língua.

Esta tese, apesar de destacar os padrões entoacionais presentes no PB, se concentra na relação dinâmica que a entoação estabelece com a fala. Na produção da fala a entoação se manifesta como a forma mais produtiva de expressão das intenções do falante. Através da produção e manipulação físicas de f_0 é possível atribuir ênfase a um determinado item, determinar se o que está sendo dito é uma asserção, uma interrogação, uma interrogação parcial, determinar foco, pressuposição e status da informação, entre outras funções.

1.2 O programa dinâmico

A perspectiva dinâmica nas ciências teve seu início a partir dos estudos da motricidade humana e animal. As pesquisas de Kugler & Turvey (1987), Thelen & Smith (1994) e Kelso (1994; 1995) sobre motricidade em geral, revelaram a existência de uma coordenação funcional e auto-organizada de movimentos que pressupõe o abandono das dicotomias mente-corpo e planejamento-execução em detrimento da existência de sistemas dinâmicos atuando no controle desses movimentos.

Um exemplo típico de coordenação entre movimentos e sistemas sensoriais é o movimento executado pelos gatos para saltar de uma árvore ou sobre alguma plataforma. Da mesma forma, a capacidade humana de saltar, andar e aprender a falar são exemplos dessa coordenação. Ao observarmos a figura 1 a seguir, podemos acompanhar os sucessivos movimentos que um gato¹ executa para saltar e cair seguramente sobre as quatro patas. Os movimentos registrados pelo fotógrafo Eadweard Muybridge² em sua clássica

¹ Fotografia estroboscópica de Kim Taylor.

² Fotógrafo inglês nascido em 1830, conhecido por seus experimentos fotográficos utilizando múltiplas câmeras a fim de captar a movimentação de homens e animais em diferentes situações. Seu trabalho é considerado precursor na criação da película de celulóide usada no cinema.

obra *Animal Locomotion* (1887) também são exemplos da organização envolvida na movimentação dos homens e dos animais em um espaço de tempo, como se observa na figura 2.



Figura 1: gato saltando; durante o salto o gato ajusta seu movimento para atingir a tarefa de cair em pé.

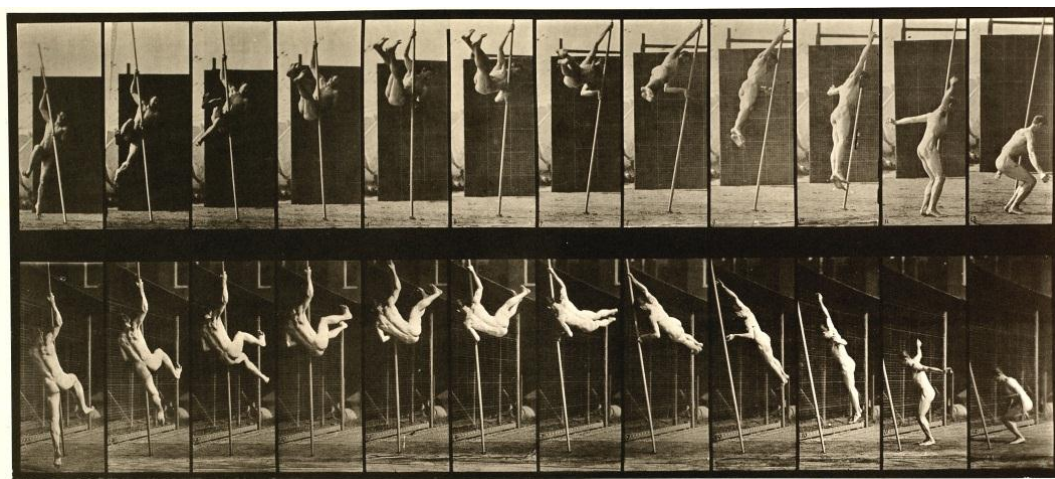


Figura 2: (*Animal Locomotion. Plate 165*) da mesma forma que o gato, o homem coordena seus movimentos para a realização de uma tarefa específica.

Essa execução perfeita de movimentos no espaço é possível pela ação de um sistema dinâmico, que compreende um sistema físico composto por um número de diferentes componentes que interagem entre si e com o ambiente de uma forma auto-organizada, e que muda seu estado inicial com a passagem do tempo (van Gelder & Port, 1995; Barbosa, 2006).

O conceito de auto-organização se refere a uma tendência espontânea não-linear de ordenamento observada em sistemas complexos, que podem ser naturais ou artificiais. Esse ordenamento, ou interação, deve-se a características intrínsecas ao próprio processo desses sistemas complexos. Tais sistemas são sensíveis às suas condições iniciais de partida, podendo estar imersos em um regime caótico³, e à sua relação com o meio ambiente.

Exemplos de auto-organização podem ser encontrados na natureza na forma de sistemas químicos, físicos, biológicos e sociais, em certos tipos de magnetismo, na estrutura de um furacão, na organização de colmeias, de movimentos sociais, na coloração de espécies animais e nos mecanismos de produção da fala.

A relação de auto-organização pressupõe o abandono das dicotomias entre mente e cérebro e planejamento e execução em detrimento de ações coordenadas. No entanto, para o controle dessas ações o número de variáveis a serem controladas devem ser reduzidas e se correlacionarem por meio de uma sinergia⁴ (Kelso, 1995), que permite que o sistema exiba propriedade de adaptação. Dessa relação resulta um padrão de funcionamento que age de acordo com uma frequência interna e que é capaz de alterar seu estado inicial ao longo do tempo. Esse comportamento possibilita ao observador a capacidade de previsão de eventos futuros a partir da observação do estado inicial desse sistema e de sua frequência em um dado momento, como por exemplo, os saltos do gato e do homem nas figuras 1 e 2.

³ É importante notar que nem todo sistema está imerso no caos. O caos é um dos regimes possíveis em alguns sistemas a partir de certos valores e parâmetros.

⁴ Sinergia (do grego *synergía*: *syn*, cooperação + *érgon*, trabalho) é o efeito ativo e retroativo do trabalho ou esforço coordenado de vários subsistemas na realização de uma tarefa complexa. O efeito resultante da sinergia entre de vários agentes que atuam de forma coordenada para um objetivo comum tem valor superior ao da soma do trabalho desses agentes efetuados separadamente.

A coordenação entre as diferentes partes desse sistema dinâmico auto-organizado possibilita a emergência de parâmetros de controle que podem influenciar o sistema como um todo. Quando ocorre algum tipo de modificação, mesmo que suave, em algum dos parâmetros de controle desse sistema, pode ocorrer em seguida o que chamamos de bifurcação no sistema. Em termos de sistemas dinâmicos, uma bifurcação é uma transição brusca em relação ao estado anterior em que se apresentava o sistema. Transições desse tipo podem ser de origem intrínseca ou ocorrer a partir de uma perturbação imposta externamente ao sistema. Quando há uma perturbação no sistema é que se torna possível determinar a relação funcional entre os parâmetros de controle, os quais podem ser identificados nesses momentos de transição (Kelso, 1995).

Por apresentar uma relação estreita com estados temporais, um sistema dinâmico, além de ter bifurcações como propriedades que ocorrem ao longo do tempo, pode dispor também de atratores cíclicos (caso seja um sistema periódico) que são padrões de movimentos oscilatórios para os quais o sistema tende com o passar do tempo ou com a realização de uma tarefa específica, ou ainda, quando sujeito a alguma perturbação externa (Barbosa, 2006).

Segundo essa perspectiva dinâmica, além da coordenação de movimentos que possibilitam a um sujeito a realização de uma tarefa específica, existe ainda uma tendência biológica à coordenação de movimentos entre sujeitos ou entre sujeito e ambiente, fato que pode ser exemplificado por dois experimentos: o de Kelso (1984), em que sujeitos ao executarem movimentos oscilatórios com os dedos indicadores das mãos em anti-fase, entram em fase quando alcançam uma velocidade alta, passando por um limiar de transição de fase; e pelo experimento reportado em Barbosa, et al (2005) sobre uma sequência de repetições de sílabas CV (consoante-vogal) do PB que entram em fase com o ritmo de um metrônomo, coincidindo os *onsets* vocálicos das sílabas com as batidas do metrônomo.

Se esta coordenação possibilita um mecanismo de controle do movimento como abordado até o momento, na produção da fala, mais especificamente da entoação, esta hipótese também pode ser aplicada, pois esta sincronia entre elementos auto-organizados pode forçar os movimentos articulatorios que possibilitam o alcance de um alvo específico na curva entoacional a coincidirem com certos ciclos articulatorios recorrentes. Existem

evidências, por exemplo, de que a sílaba sirva como estrutura coordenativa com a qual vários movimentos articulatórios estão alinhados (Fujimura, 2000). Segundo Xu (2005), isso significa que pode ser difícil para um falante manter uma relação de fase entre movimentos de f0 e sílaba senão por uma sincronia que revela um mecanismo atrator.

1.2.1 O Programa de Investigação Científica segundo Lakatos

A abordagem dinâmica, que será explorada ao longo desta tese, se insere em um programa de investigação científica dinâmico, de acordo com a definição de Lakatos (1978).

Diferentemente das propostas mais conhecidas no estudo da filosofia da ciência, Lakatos propõe um método investigativo que não pressupõe a substituição de uma teoria por outra baseando-se na falseabilidade dos pressupostos da primeira, como defendia Popper (1996), tampouco se baseia nos conceitos de degenerescência e revoluções científicas, como defendia Kuhn (1987). Lakatos (1978) propõe que as teorias vigentes em um determinado período se agrupem em um programa de investigação científica (doravante PIC).

Um PIC, segundo Silva (1998, p. 111) “(...) *não é uma hipótese isolada, nem uma mera conjunção de hipóteses: é uma série estruturada de teorias em desenvolvimento e de regras metodológicas que traçam orientações estratégicas para investigação*”. Essas regras metodológicas consistem em:

- i) um *núcleo (hardcore)*, que contém as leis explanatórias e as assunções básicas desse programa; esse núcleo vai se constituindo progressivamente, é tenazmente protegido de qualquer ataque e irrefutável por decisão metodológica de seus proponentes;
- ii) um *cinturão de proteção* do núcleo duro, constituído por hipóteses auxiliares que suportam as conjecturas fundamentais do núcleo; esse *cinturão* suporta a pressão dos testes a que o programa é submetido, por meio de reajustes que

permitam anular tentativas de falsificação e revertê-las em sucessos explicativos (Silva, 1998);

- iii) uma *heurística negativa*, que consiste em princípios metodológicos que tem função de proteger o núcleo de refutações; indica as vias de investigação a evitar por serem inconsistentes com o programa de investigação;
- iv) uma *heurística positiva*, que aponta para o desenvolvimento da cintura de proteção do núcleo; não se trata de uma heurística defensiva, mas uma via de melhoramento da cintura de proteção, fornecendo estratégias para antecipar refutações e para incorporá-las.

Lakatos defende que a ciência mantém-se crítica e em evolução pela competição de PICs distintos, que se alternam em momentos de hegemonia. Segundo sua perspectiva, a história da ciência deve ser uma história de programas de pesquisa competitivos, que se alternam hegemonicamente, e não uma sucessão de períodos de ciência normal (cf. Kuhn, 1987). Para Lakatos, a coexistência de PICs distintos numa mesma época se deve ao fato de que a falsificação de um núcleo teórico não provoca necessariamente a degeneração de um programa, pois um PIC não é abandonado caso seu núcleo seja falseado, como defende Popper (1996). Essa tenacidade dos programas faz com que seus proponentes não os abandonem em razão de um ataque ou da emergência de outro programa, mas procurem formas de manter a sua verificabilidade e seu poder de previsão mesmo que este perca sua hegemonia para outro programa.

Segundo a proposta de Lakatos (1978) sobre o que determina um PIC, podemos considerar que as teorias dinâmicas sobre motricidade em diferentes níveis partilham de um mesmo núcleo, e formam um *programa dinâmico de investigação científica* (doravante PD).

O PD, ao explorar as ações de sistemas coordenados para um determinado fim, assume uma abordagem matemática na explicação dessas relações (Saltzman, 1995) e propõe a quebra da dicotomia mente-corpo. Esse programa co-existe e se mantém em competição com um programa que propõe uma abordagem computacional dos processos

cognitivos, separando mente – sistema processador, ou *software* – e corpo – sistema executor, ou *hardware*.

O programa computacional tem como pressuposto a representação de estruturas estáticas de símbolos discretos, que são processadas sequencialmente e ciclicamente em um conjunto de módulos específicos. Esses módulos recebem da periferia do sistema estímulos sensoriais em forma de representações – *inputs* – e geram representações simbólicas – *outputs* –, que se convertem em estímulos físicos na periferia do sistema.

Outro fator importante na distinção entre esses dois programas é a relação temporal. Enquanto um sistema dinâmico possui a propriedade de se desenvolver e se modificar ao longo do tempo real, um sistema computacional trabalha com uma sequência interna de estados em um “passo a passo” arbitrário (van Gelder & Port, 1995) que demanda tempo, porém não se desenvolve no tempo real. Nas palavras de van Gelder e Port (1995), o programa cognitivo computacional “*leave time out of the picture, replacing it only with ersatz “time”: a bare, abstract sequence of symbolic states*” (van Gelder & Port, 1995, p.3).

Na defesa de um PD, esses autores afirmam que:

The cognitive system is not a computer, it is a dynamical system. It is not the brain, inner and encapsulated; rather, it is the whole system comprised of nervous system, body, and environment. The cognitive system is not a discrete sequential manipulator of static representational structures; rather, it is a structure of mutually and simultaneously influencing change (van Gelder & Port, 1995, p.3).

e ainda:

Dynamical models are increasingly prominent in cognitive psychology, developmental psychology, and even some areas of linguistics. In short, the dynamical approach is

not just some new kid on the block; rather, to see that there is a dynamical approach is to see a new way of conceptually reorganizing cognitive science as it is currently practiced (van Gelder & Port, 1995, p.4).

Na Linguística Geral, a tentativa de analisar o desenvolvimento científico sob a ótica dos PICs de Lakatos é rara (Borges, 2004). No entanto, os programas cognitivos dinâmico e computacional podem partilhar seus núcleos com teorias linguísticas.

A teoria gerativa de Chomsky partilha dos pressupostos do programa cognitivo computacional, na medida em que considera, numa lógica cartesiana, a separação entre mente e corpo e o processamento modular da linguagem. Chomsky já propusera em seu trabalho inicial uma ruptura com o estruturalismo e o behaviorismo e o estabelecimento do que podemos chamar de um novo PIC na linguística a partir de suas ideias sobre o inatismo da linguagem⁵. Porém, observando o progresso de sua teoria e a assimilação de novos conteúdos a esta por meio de uma heurística positiva, podemos observar a teoria gerativa como integrante de um PIC maior, o PIC cognitivo computacional.

Esse PIC se mostra em hegemonia nos estudos linguísticos, em especial nas áreas da sintaxe, semântica, morfologia e fonologia. No entanto, nos estudos da fonética, em subáreas como articulação - produção, entoação e ritmo, o PIC dinâmico se encontra em concorrência com o PIC computacional e em crescente destaque. O modelo escolhido nesta tese para a análise dos fenômenos entoacionais e rítmicos envolvidos na fala, como dito anteriormente, é um modelo dinâmico.

⁵ Para uma discussão maior sobre filosofia da ciência e o programa de pesquisa Chomskyano ver Pires de Oliveira (2010).

1.3 Modelo Dinâmico do Ritmo

Dentro do PD, o modelo que se ajusta à análise entoacional desta tese é o modelo dinâmico do ritmo da fala (doravante MDR). O MDR, proposto por Barbosa (2006) é uma implementação matemático-computacional de um sistema dinâmico do ritmo da fala que exhibe todas as propriedades de um sistema auto-organizado (Barbosa, 2006) como (i) mudanças de estado dos sistemas com o passar do tempo, (ii) previsões sobre mudanças de estado em momentos futuros, (iii) sinergia e acoplamento funcional entre os componentes do sistema, e (iv) a presença de atratores cíclicos.

De acordo com esses pressupostos, o MDR recebe como entrada informação sobre níveis linguísticos elevados (sintaxe, semântica, discurso) e também um léxico gestual, segundo a definição da Fonologia Articulatória⁶ (Browman & Goldstein, 1986), e gera como saída a duração gestual (Barbosa, 2006). Ou seja, o modelo em si é composto pela (i) interação entre a informação linguística contida nos níveis superiores e um oscilador acentual, (ii) pelo acoplamento de dois osciladores linguísticos abstratos (acentual e silábico) regidos por uma força de acoplamento, (iii) e por um mecanismo de acoplamento entre o oscilador silábico e a pauta gestual⁷ (originada do léxico gestual). A interação entre oscilador silábico e pauta gestual gera a duração gestual. A figura 3 mostra o diagrama do MDR.

Nesse modelo, o oscilador acentual recebe informações especificadas de forma abstrata no léxico gestual, que determina a posição dos acentos lexicais e faz com que os pulsos do oscilador acentual ocorram nas sílabas marcadas lexicalmente como tônicas (Barbosa, 2006). Essa sucessão de pulsos especifica os locais onde ocorrem os acentos

⁶ A Fonologia Articulatória representa as unidades linguísticas em termos de movimentos articulatorios coordenados, auto-organizados e abstratos, que determinam os gestos. Os gestos articulatorios, como unidades de representação, podem explicar processos fonológicos, defendendo a ideia de que todos esses podem ser entendidos como resultantes da sobreposição e da magnitude dos gestos.

⁷ As pautas gestuais são uma representação abstrata do planejamento do enunciado e estão relacionadas aos movimentos do trato vocal, que, por sua vez, são representados como descontínuos. As pautas são formadas dispondo os gestos em camadas verticais, as quais são definidas usando a noção de independência articulatória, e cada gesto é representado por retângulos de base menos ou mais longa, de acordo com a duração abstrata, além de menos ou mais altos, de acordo com a magnitude.

frasais isomorficamente, ou seja, obedecendo a um mesmo padrão. Dessa forma, o oscilador acentual gatilha no oscilador silábico um mecanismo de acentuação universal ao longo dos ciclos do oscilador silábico, que gera os padrões de duração de unidades V-V (vogal-a-vogal)⁸ ao longo do enunciado, assinalando os acentos frasais (Barbosa, 2006).

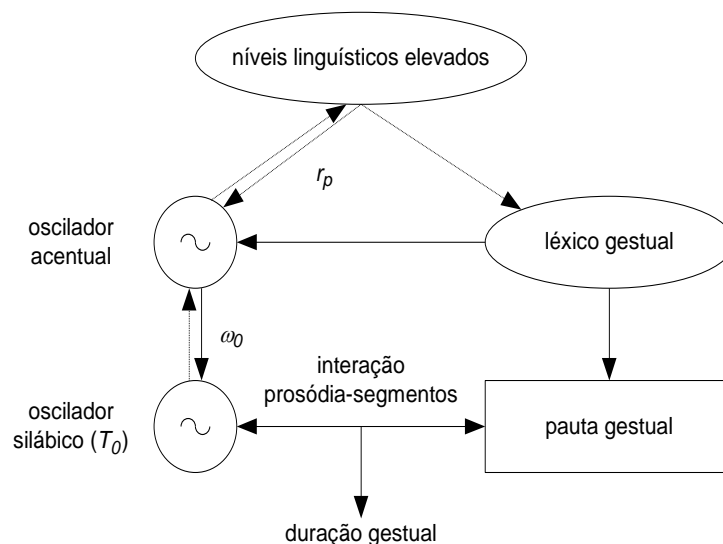


Figura 3: Diagrama do Modelo Dinâmico do Ritmo da Fala

O acento frasal, definido por um pico de duração local no enunciado, assinala uma proeminência do domínio da produção da fala, baseada apenas na duração dos segmentos. A interação entre essa proeminência, enquanto um fator prosódico, e a pauta gestual, gera a duração gestual. Barbosa (2006, p.11) destaca que a ênfase entoacional, obtida com a excursão de f_0 em determinado local, acarreta aumento de duração nos segmentos vocálicos, e é mais saliente perceptivelmente do que apenas o aumento das durações. Esse alinhamento entre f_0 e duração dá pistas sobre como o acento frasal pode

⁸ As unidades V-V compreendem o espaço entre os *onsets* de duas vogais consecutivas. Uma explicação mais detalhada sobre o assunto está no capítulo Notação e Descrição Entoacional.

influenciar não somente a estrutura rítmica dos enunciados do PB, mas também a estrutura entoacional.

1.4 Sumário

- a entoação é definida por características fonéticas suprasegmentais que expressam significados pragmáticos de forma estruturada nos enunciados; a entoação se liga à fala, numa estrutura particular, e à língua, determinando seus padrões entoacionais;
- a estrutura de movimentos no espaço, assim como os movimentos envolvidos na produção de f_0 , são possíveis pela ação de um sistema dinâmico;
- um sistema dinâmico compreende um sistema físico formado por diferentes componentes que interagem entre si e com o ambiente de forma auto-organizada;
- existe um PIC dinâmico que se contrapõe a um PIC computacional hegemônico nos estudos da linguagem;
- o MDR se insere no PIC dinâmico e propõe a ação de dois osciladores acoplados (acentual e silábico) para explicar a produção do ritmo da fala;
- esse modelo pode ser adotado para representar também a estrutura subjacente da entoação.

II FALA E DISCURSO

1. CORPUS DE FALA

O corpus de fala utilizado nesta pesquisa foi composto com o intuito de trabalhar com um conjunto de dados que pudesse refletir a dinâmica própria do funcionamento da língua falada no PB, ou seja, a dinâmica dos usos formais e informais, interação entre falantes e espontaneidade. Seguindo esse ideal, foi montado o *Corpus VoCE*, acrônimo formado pelos nomes da rádio **V**ocê AM, do site **C**iência Hoje *O*nline, e do programa de rádio Fim de **E**xpediente dos quais foram obtidas as gravações diretas dos programas de rádio e de *podcasts*⁹.

O Corpus VoCE está em composição desde 2006 (Lucente e Barbosa, 2007) e conta com aproximadamente 10 horas de gravações no total, contendo exemplos da fala de profissionais (jornalistas ou radialistas) e de participantes dos programas. Os 46 trechos selecionados para análise a partir do corpus somam aproximadamente 32 minutos de duração e são exemplos de fala de 23 sujeitos diferentes, sendo estes nove mulheres e 14 homens. Desses exemplos foram excluídas as falas dos profissionais, por apresentarem registro com marcas de locução profissional, que não é o foco do presente trabalho.

Os trechos de fala selecionados apresentam em sua maioria exemplos da fala paulista e carioca, pois as rádios e a revista de onde foram extraídas as gravações estão localizadas em São Paulo (Rádio Você AM, Americana - SP; Programa Fim de Expediente gravado em São Paulo, capital) e Rio de Janeiro (Revista Ciência Hoje *O*n Line, produto da Fundação Oswaldo Cruz, Rio de Janeiro, capital). Informações sobre os sujeitos, como sexo, faixa etária e escolaridade foram obtidas a partir do reconhecimento dos falantes, que são em grande parte figuras públicas. Tais informações, assim como as durações dos trechos de falas analisados estão apresentadas na tabela 1 a seguir.

A estrutura e a composição do Corpus VoCE se assemelham às de outros corpora compostos por gravações radiofônicas, como por exemplo o *Boston University Radio News Corpus* (Ostendorf et al, 1996), que foi bastante utilizado em pesquisas

⁹ Respectivamente: <http://www.radiovoce.com.br>; <http://cienciahoje.uol.com.br/podcasts>; <http://cbn.globoradio.globo.com/programas/fim-de-expediente/FIM-DE-EXPEDIENTE.htm>

fonético-entoacionais do inglês (Hasegawa-Johnson et al, 2005; Abner, 2009; Rosenberg & Hirschberg, 2009). Ambos os corpora classificam seus dados como sendo exemplos de fala espontânea, devido ao fato de estes não serem obtidos em laboratório. No entanto, tal classificação dos dados de fala como sendo espontâneos merece especial atenção, como se verá a seguir.

Sujeito	Idade	Sexo	Educação	Dialeto	Dur.(s)
AA	20-30	F	ensino médio	paulista	81,19
CL	40-50	F	superior	paulista	24,16
CM	30-40	F	ensino médio	carioca	19,37
CP	40-50	F	superior	carioca	154,27
DS	40-50	M	superior	paulista	16,01
DS	30-40	M	ensino médio	alagoano	17,88
GK	50-60	F	superior	paulista	225,13
JG	40-50	M	superior	paulista	141,41
JG(2)	20-30	M	ensino médio	paulista	30,73
JL	40-50	F	técnico	gaúcha	47,53
JO	30-40	M	superior	carioca	145,64
JR	40-50	M	superior	carioca	283,98
LB	40-50	M	superior	carioca	63,65
LF	40-50	F	superior	paulista	106,61
LR	50-60	F	superior	carioca	71,18
MA	40-50	M	superior	carioca	32,48
MG	40-50	M	superior	mineiro	49,91
MP	30-40	M	superior	paulista	45,31
MS	30-40	M	superior	carioca	71,91
MS(2)	30-40	M	ensino médio	brasiliense	24,19
RI	50-60	M	superior	carioca	60,23
RV	50-60	F	ensino médio	paulista	130,11
VA	50-60	M	superior	carioca	98,67

Tabela 1: resumos das propriedades do conjunto de sujeitos do corpus.

1.1 Em defesa da fala espontânea

A pesquisa com fala espontânea na fonética é um método cercado por discussões empíricas acerca de sua taxonomia e forma de análise, tendo inclusive algumas publicações dado especial atenção ao tema, como em Sagisaka, et al (1996). Na introdução à primeira parte desta edição, Ladd (1996b) comenta que talvez a única forma de investigar o que faz da fala natural seria a desconstrução de sua naturalidade por meio da experimentação.

Nos estudos sobre entoação, a pesquisa com dados de fala espontânea da forma como é feita atualmente obteve grande impulso nos anos 1990 e teve como destaque trabalhos como os de Hirschberg & Litman (1993), Nakatani & Hirschberg (1994), Ostendorf et al (1996), Beckman (1996), Nakatani (1996), Campbell (1996), Kholer (1996). A pesquisa com esse tipo de dado é também explorada no PB, como se vê em Lucente (2008), Lucente & Barbosa (2007, 2008, 2010), Barbosa (2008, 2010), Côrtes et al (2011) e na composição do corpus oral C-Oral Brasil (Raso & Mello, 2012).

Em sentido oposto dessa tendência na pesquisa fonética, os trabalhos de Xu (1999, 2005, 2006), em especial seu artigo *In defense of lab speech* (2010), têm mostrado que ao contrário do que se pregava com o uso da fala espontânea nas pesquisas, a fala de laboratório, ou *lab speech*, se mostra bastante produtiva para análise de fenômenos fonético-entoacionais.

Xu (2010) levanta diversos pontos em defesa da fala de laboratório, partindo de um senso comum bastante compartilhado na comunidade fonética de que somente a conversa espontânea pode mostrar a natureza da fala. Segundo o autor, essa visão é acompanhada da ideia de que a fala de laboratório é insuficiente para representar a riqueza da fala espontânea. Para mostrar que essa ideia é equivocada ele procura desconstruir mitos em torno do uso da fala de laboratório na pesquisa linguística, como por exemplo, o fato de a fala de laboratório ser lenta, pouco natural, muito planejada, monótona em relação à prosódia, desprovida de funções comunicativas, interativas, e de emoções.

Para defender a ideia de que a fala de laboratório não possui tais características, Xu (2010) usa como argumento o fato de que toda fala é natural por definição, desde que não seja produzida por síntese. A fala em qualquer circunstância é um fenômeno natural, e pode ser produzida em diferentes taxas de elocução, de acordo com a orientação do experimentador, o que derruba o mito de que a fala de laboratório não seria natural ou seria lenta demais. Sobre o fato de a fala de laboratório ser super planejada, o problema se resume à forma como são apresentados os estímulos aos sujeitos, o que mostra a importância do experimentador em evitar uma situação em que o sujeito do experimento tenha meios para planejar demasiadamente sua fala. Ou seja, o autor enfatiza que a fala de laboratório é capaz de apresentar os mesmos recursos atribuídos à fala espontânea desde que o controle do experimento seja executado a fim de obter fenômenos específicos na fala dos sujeitos.

A definição de fala espontânea e não espontânea proposta por Beckman (1996), e também adotada por Xu (2010), sugere que os dados de fala espontânea são aqueles obtidos fora das condições de laboratório, sem qualquer intervenção de um experimentador, enquanto que os dados obtidos em laboratório por meio de leitura e outros tipos de eliciações, mesmo que levem a um grau de espontaneidade, são considerados fala de laboratório, ou *lab speech*. Nesse sentido, vale lembrar que a grande maioria dos exemplos a serem considerados aqui são de fala obtida livremente, fora de um laboratório ou com propósitos de pesquisa, ou seja, na concepção de Xu e Beckman, fala espontânea.

A fala espontânea apresenta vantagens nos dados sobre entoação, pois, se por um lado, a pesquisa fonético-articulatória segmental utiliza medidas como duração, intensidade, articulação, a pesquisa com entoação, por outro lado, requer a medida da frequência fundamental, que é mais robusta e adequada em trechos com menor relação sinal-ruído¹⁰, como os exemplos obtidos para o Corpus VoCE, sobretudo os exemplos extraídos do *podcast* do programa “Fim de Expediente”, que por vezes é gravado em auditório com plateia. Outra argumentação importante na defesa da fala espontânea é a de

¹⁰ Relação sinal-ruído (SR), ou *signal to noise ratio* (SN), é a relação entre a potência de um sinal (música, fala) e a potência do ruído sobreposto a esse sinal. Quanto mais alto o valor dessa relação, menor é o efeito do ruído de fundo sobre a detecção ou medição do sinal.

que a eliciação de padrões entoacionais específicos em laboratório é complexa, e pode resultar em dados ambíguos, como os descritos em Lucente e Barbosa (2007). Nesse artigo, dentre vários experimentos, é descrita a tentativa sem sucesso de eliciação de sentenças exclamativas. A forma como foi conduzida a gravação levou os sujeitos a produzirem declarativas neutras nos contextos onde se esperavam exclamações.

Outro aspecto importante a se considerar na defesa da fala espontânea é que esta é bastante produtiva e de encontro aos interesses de análises que levam em conta a conversação, pois as reais intenções de um falante, que carregam um conteúdo genuinamente pragmático, podem ser observadas claramente em um trecho de fala sem qualquer controle experimental. A fala espontânea apresenta riqueza de estilos e dialetos que podem facilmente ser obtidos em laboratório, no entanto nem sempre ocorre o mesmo com o léxico, pois da fala espontânea podem surgir palavras e expressões inesperadas para o experimentador.

Podemos contra argumentar ainda sobre as vantagens da fala espontânea assumindo um ponto na exposição de Xu (2010), sobre o fato de que em outras ciências a experimentação raramente é feita com dados reais e/ou naturais, mas sim com amostras e situações criadas em laboratório, como ocorre na química, física e biologia, por exemplo. Deixando de lado o caráter positivista presente nessa comparação entre a forma de experimentação na linguística e nas ciências exatas e biológicas, é importante considerar que a pesquisa linguística trata de material humano e em constante modificação. Sendo assim, a possibilidade de lidar com dados extraídos da situação de uso podem ser encarados como uma vantagem sobre as ciências que não têm a possibilidade de obter dados dessa espécie para análise.

Contudo, para se adquirir habilidade da análise dos dados de fala espontânea, a experimentação com dados de laboratório é essencial em algum momento da pesquisa. O recorte no objeto de estudo e a composição de um corpus experimental obtido em laboratório é importante para se observar o objeto de pesquisa, o comportamento dos dados e a direção para qual apontam as descobertas (Lucente, 2008). De posse desse material e experiência, a observação e a análise dos dados de fala espontânea ocorrem de forma mais direcionada. Por exemplo, a pesquisa que deu origem a essa tese se iniciou a partir de um

corpus obtido em laboratório em que os falantes foram induzidos a produzir enunciados em diferentes modalidades e com foco em diferentes partes do enunciado. Dessa forma a análise dos padrões entoacionais no PB pôde começar a ser delineada. Sendo assim, o passo seguinte da pesquisa foi observar se os mesmos padrões estavam presentes na fala espontânea e com qual variabilidade (Lucente, 2008). Sobre essa necessidade de o linguista saber lidar com dados de laboratório, mesmo que este esteja interessado em dados espontâneos, Ladd (1996b) comenta:

... linguists can and should work in the laboratory, and that the theoretical constructs of linguistics are in principle relevant to describing the behavior of actual speakers. This point has sometimes been forgotten, both by linguists – whose academic culture to a great extent prizes theorizing over applications – and by experimental phoneticians and engineers, who have often tended to belittle, dismiss, or ignore what linguists have to say. This has been especially true in the area of prosody, where a real split between “experimentalist” and “impressionistic” approaches was evident especially during the period from about 1950 to 1980. (Ladd, 1996b, p.4)

Feita a defesa da fala espontânea e explorada a dicotomia entre esta e fala de laboratório, surge ainda um ponto importante a ser discutido: o que define a fala como sendo espontânea? A definição adotada de que a fala espontânea é a fala que não é lida a partir de um roteiro versus a fala de laboratório, que é obtida sob controle experimental (Xu, 2010; Beckman, 1996), nem sempre é satisfatória para tal classificação. O limite no qual a fala se distingue entre espontânea ou de laboratório é bastante tênue, a classificação do dado como espontâneo acaba por ser feita pelo pesquisador, classificação que por vezes acaba não sendo clara ou imparcial, visto que são comuns as discussões entre pesquisadores e audiência entre o que cada parte considera como sendo dado espontâneo.

Refletindo sobre esse impasse e sobre a necessidade de uma classificação mais efetiva para os dados de fala é proposto, portanto, um critério para classificar os dados

como sendo de fala espontânea ou não espontânea. Esse método propõe a classificação dos dados de fala de acordo com um grau de espontaneidade definido entre o cruzamento de dois parâmetros: i) o gênero no qual o discurso de encaixa; e ii) o controle externo sobre a produção do discurso.

1.2 Classificação da fala espontânea: gênero e controle

As definições adotadas por Xu (2010) e Beckman (1996) deixam de abordar, ou abordam sem desenvolver suas implicações, dois pontos essenciais para a definição da fala espontânea, que são (i) o grau de controle do experimentador sobre a obtenção dos dados, e (ii) o papel dos gêneros discursivos .

A fala espontânea não pode ser considerada sinônimo de conversa livre apenas, como já apontado, pois um ditado, uma leitura, uma narrativa ou uma entrevista podem ser considerados espontâneos se ocorrerem em situação natural de comunicação. Da mesma forma, fala de laboratório não é sinônimo de total ausência de espontaneidade, pois dependendo da forma como os sujeitos são eliciados estes podem alcançar bastante espontaneidade na fala. O que irá definir a dimensão que caracteriza a espontaneidade é o grau de intervenção do experimentador na obtenção dos dados. Essa perspectiva de que graus de controle podem determinar a naturalidade da fala já foi abordada por Lehist (1963):

(...) the linguist who wants to use experimental methods in his investigation is forced to trade a certain amount of naturalness for rigorousness in his experimental design (...) The two requirements of naturalness and rigorousness appear to occupy the opposite ends of a continuous scale of gradual mutual exclusion. At one extreme we find complete naturalness coupled with complete lack of control; at other end, complete control over experimental design at the expense of naturalness and linguistic relevance. (apud, Ladd, 1996b, p. 3).

Sendo assim, a intervenção, ou controle do experimentador sobre a obtenção dos dados vai de um controle extremo – situação em que o experimentador diz como o sujeito deve produzir a fala – até a ausência total de controle – fala dita sem nenhuma indicação de produção, em qualquer que seja seu registro¹¹. De acordo com Barbosa (2012), o termo espontâneo usado restritivamente para uma conversa, perde seu sentido, pois a fala espontânea pode se dar em diversos gêneros discursivos.

A proposta de classificação por meio da relação entre dois eixos leva à necessidade de se descrever o gênero discursivo que se pretende classificar. Os gêneros discursivos identificados no Corpus VoCE são fala espontânea, entrevista, aula/palestra, leitura, narração e gênero jornalístico. Tais gêneros podem exibir diferentes níveis de controle do experimentador sobre a eliciação, se houver eliciação. Dessa forma, se faz necessário a descrição do gênero de fala e da situação de controle sobre sua obtenção para que se possa classificar efetivamente um corpus de fala como espontâneo, semi-espontâneo ou não espontâneo.

Aplicando esse método ao Corpus VoCE, a partir de uma relação ortogonal entre os fatores gênero e controle, foi possível obter uma classificação.

1.2.1 Classificação do corpus

De acordo com a classificação proposta, no eixo do gênero, estão classificados os gêneros discursivos de forma não ordenada, pois não é possível ordená-los segundo uma hierarquia em que um gênero apresente mais naturalidade em relação a outro (Barbosa, 2012), pois cada um pode ser realizado de acordo com uma escala de naturalidade própria a seu gênero.

A escala de naturalidade se realiza no eixo do controle, que classifica o grau de intervenção de um experimentador na obtenção dos dados, sendo as extremidades desse eixo marcadas pela ausência completa de controle e pelo controle extremo do

¹¹ De acordo com uma perspectiva sociolinguística dos gêneros e registros de fala (não textual), Bieber & Conrad (2009) defendem que o registro é uma variedade associada a uma situação particular de uso, que inclui propósitos particulares. A descrição do registro nesta perspectiva cobre três componentes maiores, que são: o contexto; as características linguísticas; e as relações funcionais entre os dois primeiros componentes.

experimentador. Seguindo esse eixo, a produção da fala pode ter níveis de controle distintos exercidos por um experimentador. Por exemplo, em uma entrevista, o controle sobre a produção da fala pode ser feito pelo próprio experimentador-entrevistador, que é quem conduz a entrevista. Este pode conduzir a entrevista a um controle extremo em que o entrevistador restringe as respostas do entrevistado a apenas um assunto ou formas de resposta (sim ou não, por exemplo), ou casos em que entrevista não possui nenhuma forma de controle, e o entrevistado é livre para responder as perguntas da forma que preferir. A classificação dos gêneros discursivos em relação ortogonal ao controle exercido em sua obtenção por ser observada na figura 4.

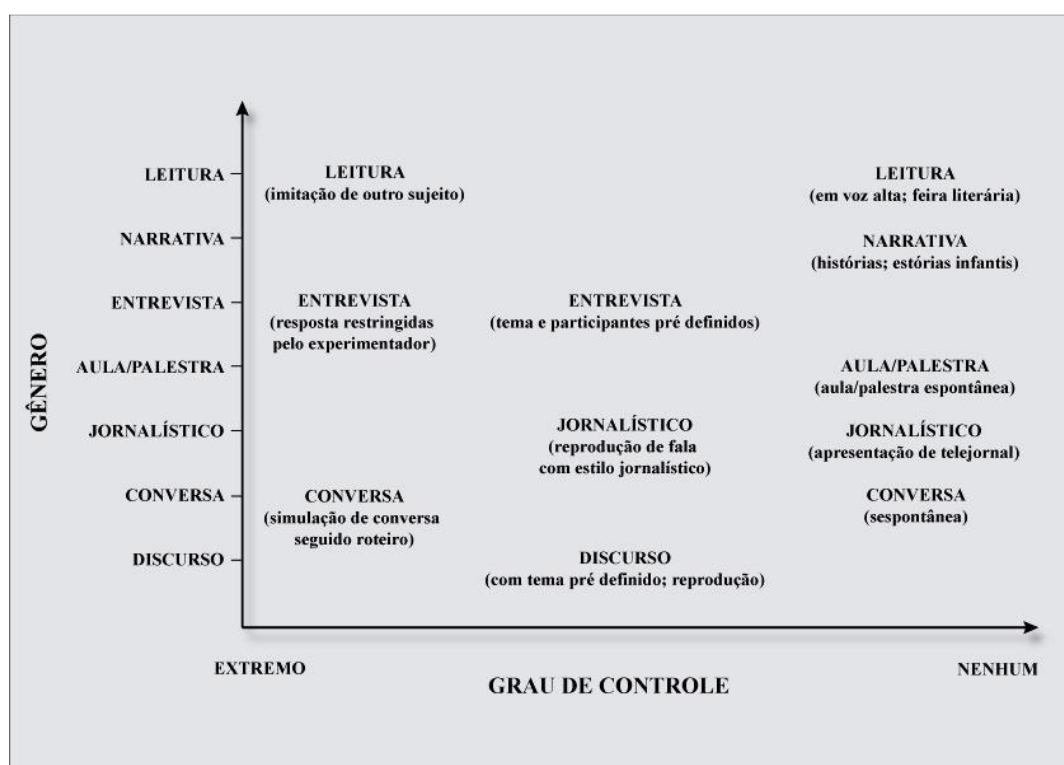


Figura 4: classificação dos gêneros discursivos em relação ao eixo que mede o grau de controle.

De acordo com essa classificação, o gênero leitura, por exemplo, que é comumente tido como não espontâneo, pode ser classificado como espontâneo se a leitura em questão for realizada em um evento, uma feira literária, etc. Ao contrário, se a leitura

tiver como objetivo fazer uma imitação da forma como outra pessoa lê, é classificada como não espontânea.

Essa proposta de classificação contribui para evitar interpretações diferentes acerca da espontaneidade da fala se baseando apenas nos gêneros e não considerando a intervenção externa no momento de sua produção.

1.2.2 Teste de Percepção

Para verificar como os ouvintes classificam trechos de fala como espontânea e se tal classificação apenas de outiva se aproxima da classificação proposta nesta tese, foi realizado um teste de percepção de escolha forçada (*Multiple Forced Choice – MCF*)¹². Para a realização do teste foi utilizada a plataforma virtual *Survey Gizmo – Student Edition*¹³, que oferece recursos para realização de pesquisas e testes *online*¹⁴.

Participaram do teste 26 sujeitos, dentre os quais seis homens e 20 mulheres, com idade média de 27 anos, todos com nível superior de ensino. Após aceitarem os termos de realização, tais como utilizar fones de ouvido e informar alguns dados pessoais (nome, idade, local de nascimento e escolaridade), o teste partia para a reprodução de um conjunto de 23 estímulos sonoros não aleatorizados de 13 diferentes locutores, contendo trechos de diferentes discursos do Corpus VoCE com duração média de 10s. Para cada um dos 23 estímulos, os sujeitos deveriam classificar (i) o nível de espontaneidade na fala destes numa escala de Likert¹⁵ de 1 a 7, e (ii) associar livremente cada um dos estímulos a um gênero ou situação de fala. A figura 5 ilustra como o teste foi apresentado aos sujeitos.

¹² Testes de escolha forçada, ou múltipla escolha forçada, são testes em que os sujeitos têm necessariamente que fazer uma escolha ou atribuir uma propriedade a determinada questão.

¹³ <https://students.sgizmo.com/>

¹⁴ Um link para o teste foi publicado numa rede social (<http://facebook.com>) com um convite para que as pessoas fizessem o teste, e com as instruções de que os participantes deveriam ser brasileiros e se certificarem de que estavam com uma boa conexão de internet.

¹⁵ A escala de Likert (Rensis Likert – 1903-1981) é a escala de escolha mais comum em teste de escolha forçada. Essa escala consiste em escolhas de respostas ordenadas por números e notas, geralmente entre 1 e 5, 1 e 7 ou 1 e 9.

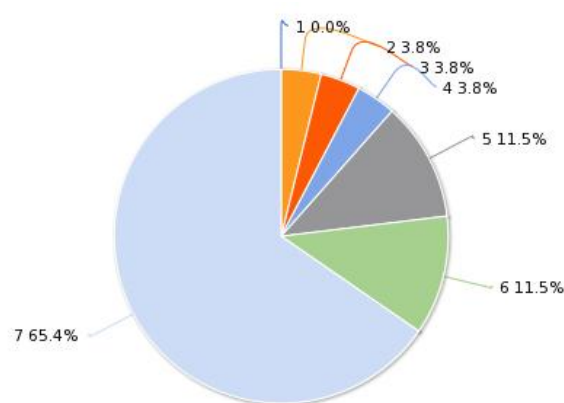
O resultado das escolhas dos sujeitos foi submetido a dois testes Kappa Fleiss¹⁶ de concordância entre sujeitos no programa R (versão 2.14.0). O primeiro teste considerou as sete opções de escolha dos sujeitos e obteve o resultado $k=0,04$, e parâmetro $z=8,39$ ($p<10^{-7}$), que mostra resultado significativo para concordância muito baixa. O segundo teste, agrupando os resultados apenas três níveis de espontaneidade (1 a 3 passaram a ser 1; 4 passou a ser 2; 4 a 6 passaram a ser 3) obteve resultado $k=0,138$ e $z=15,0$ ($p<10^{-7}$), que também mostra um resultado significativo para concordância baixa.

Figura 5: apresentação do teste MFC *online*, na plataforma Gizmo

Tais resultados mostram que a classificação da fala como espontânea pelos sujeitos não é uniforme, revelando que um gênero de fala, apesar de classificado como espontâneo no eixo de controle externo (pois é sabido que se trata de fala obtida sem controle), não parece aos sujeitos como fala espontânea. Esse é o caso da leitura, que pode ser classificada como espontânea quando não há intervenção de um experimentador sobre ela, porém os sujeitos não reconhecem esse gênero como espontâneo. Os resultados, por outro lado, podem indicar o fato de os sujeitos não serem precisos em indicar que um

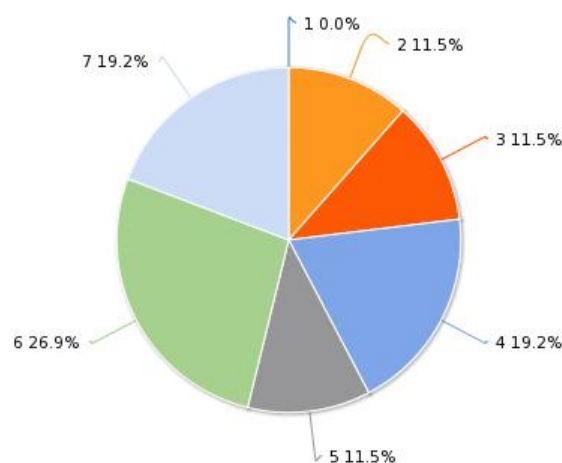
¹⁶ Kappa Fleiss é um teste estatístico de concordância entre n sujeitos, diferentemente do teste Cohen Kappa, que mede concordância entre dois sujeitos. Nos testes Kappa, quanto maior o coeficiente de concordância entre 0 e 1, maior a concordância entre os sujeitos.

discurso é espontâneo baseando-se apenas em estímulos sonoros. No gráfico da figura 6, que representa as escolhas entre 1 e 7 para um trecho da gravação “usp” (ver Anexo II), em uma ordem crescente do menos ao mais espontâneo, 65,4% dos sujeitos classificaram os exemplos como muito espontâneos. Esse seria um resultado esperado, uma vez que os exemplos apresentados foram de fala espontânea, ou seja, sem intervenção do experimentador. Porém, no gráfico da figura 7, referente à gravação “jogador-futebol”, as respostas foram bastante distribuídas entre todos os graus, desde o pouco espontâneo ao muito espontâneo, o que mostra que não houve consenso sobre a espontaneidade da fala.



Grau de espontaneidade	Nº de escolhas por grau	Porcentagem nº de escolhas
1	0	0%
2	1	3,8%
3	1	3,8%
4	1	3,8%
5	3	11,5%
6	3	11,5%
7	17	65.4%

Figura 6: gráfico e tabela de respostas para a gravação “usp”: o numero de escolhas por grau soma 26, que é o número de sujeitos; observe que 17 (65%) sujeitos atribuíram valor 7 de espontaneidade.



Grau de espontaneidade	Nº de escolhas por grau	Porcentagem nº de escolhas
1	0	0%
2	3	11,5%
3	3	11,5%
4	5	19,2%
5	3	11,5%
6	7	26,9%
7	5	19,2%

Figura 7: gráfico e tabela de respostas indicando homogeneidade na escolha entre os graus de espontaneidade para a gravação “jogadores_futebol”; observe que as escolhas por graus são muito próximas, diferentemente do gráfico anterior.

O léxico, a taxa de elocução e o registro¹⁷ dos exemplos da fala podem influenciar na atribuição de espontaneidade à fala. Isso fica evidente quando observamos os gráficos de resultados do teste (ver Anexo I), que mostram como a classificação dos estímulos não foi homogênea daí a importância de um método de classificação que indique antecipadamente o gênero e o controle dos dados.

Quanto à segunda tarefa solicitada no teste, que era classificar a fala de acordo com o gênero (estilo) que os ouvintes reconheçam, a maioria das respostas se concentrou nos seguintes gêneros: entrevista, aula/palestra, fala jornalística, narração, leitura e conversa informal. O Anexo I mostra as respostas originais dos sujeitos (inclusive com erros de grafia) que foram agrupadas em classes menores de gêneros. Por exemplo, o termo explicação foi agrupado em aula/palestra, reportagem e locução foram agrupados em

¹⁷ Na concepção sociolinguística de Biber e Conrad (2009)

gênero jornalístico. Esses gêneros apontados pelos participantes do teste foram usados para exemplificar as relações entre gênero e controle na figura 4. Ou seja, apesar do julgamento dos sujeitos não ser uniforme quanto ao grau de espontaneidade, a associação destes com um gênero específico pode ser verificada.

1.3 Sumário

- A estrutura do Corpus VoCE, composto por gravações radiofônicas, é a de exemplos de fala espontânea, devido ao fato de estes não serem obtidos em laboratório;
- Os dados de fala nem sempre são percebidos ou classificados como espontâneos, pois dentro do estudo da fala a classificação de tais dados não é consensual;
- Foi proposta, portanto, a classificação dos dados de fala de acordo com um grau de espontaneidade definido entre o cruzamento de dois parâmetros: i) o gênero no qual a fala se encaixa; e ii) o controle externo sobre a produção da fala;
- Perceptualmente os sujeitos não são precisos em indicar se um dado exemplo de fala é espontâneo baseando-se apenas em estímulos sonoros, daí a importância da classificação;
- A classificação dos gêneros de fala contidos no Corpus VoCE a partir da proposta apresentada se concentrou nos gêneros entrevista, aula/palestra, fala jornalística, narração, leitura e conversa informal;
- O grau de controle na obtenção de fala dos exemplos contidos no Corpus VoCE se concentram na extremidade da ausência de controle.

2. ESTRUTURA DO DISCURSO

A análise da fala e da entoação presentes nesta tese parte de uma abordagem discursiva que compreende as partes que compõe o discurso. Serão adotados aqui os termos “análise discursiva” e “análise do discurso”. Porém, a análise apresentada não se insere na perspectiva da Análise do Discurso enquanto área do saber. A análise discursiva, que segue uma teoria computacional sobre a estrutura discursiva (Grosz & Sidner, 1986), se aproxima também de uma perspectiva textual-interativa da conversação (Jubran & Koch, 2006), na qual se assume uma inter-relação cognitiva entre os participantes de uma conversa, tendo como objeto de estudo o texto, nesse caso, o texto falado, o qual será denominado daqui por diante por *discurso* apenas.

A análise textual-interativa da conversação, inserida na linguística textual, leva em consideração, além do aspecto cognitivo, aspectos recorrentes da subjetividade dos falantes, como a produção de sentido, por exemplo. A abordagem do discurso utilizada aqui procura explorar apenas os aspectos formais e funcionais que levam à produção do discurso e seu entendimento por parte dos interlocutores, ou seja, um processo cognitivo que se encerra quando a comunicação acontece, não sendo relevante para a análise discursiva a interpretação da produção de sentido por parte do sujeito.

2.1 Estrutura do Discurso - Modelo de Grosz e Sidner

A segmentação discursiva dos dados do Corpus VoCE foi feita de acordo com a Teoria da Estrutura do Discurso ou Modelo de Grosz e Sidner (*G&S's model* – modelo G&S em PB), proposta em Grosz & Sidner (1986). O modelo G&S foi desenvolvido como um modelo computacional que, ao descrever a estrutura do discurso, oferece as bases necessárias para sua descrição e seu significado. De acordo com as proponentes desse modelo, a descrição da estrutura do discurso desempenha um papel central no

processamento da linguagem à medida que estipula restrições nas porções do discurso (Grosz & Sidner, 1986). Essa descrição é intimamente relacionada a duas questões: *o que distingue o discurso, e o que o faz coerente?* A tentativa de resposta a essas questões leva a dois aspectos não linguísticos fundamentais no desenvolvimento desse modelo, que são a *atenção* e a *intenção* dos participantes de uma conversa. A atenção é um fator essencial no processamento de enunciados em um dado discurso, enquanto a intenção desempenha um papel importante na explicação de como se estrutura o discurso, proporcionando coerência a este e ao próprio termo “discurso” (Grosz & Sidner, 1986, p. 175).

A hipótese defendida nesse modelo é de que qualquer discurso é composto por três elementos essenciais e distintos, mas que interagem entre si a todo momento, que são: i) a estrutura sequencial dos enunciados do discurso em um dado momento; ii) a estrutura das intenções envolvidas no discurso; e iii) o estado de atenção dos participantes envolvidos no discurso. Paralelamente, existem ainda dois fatores importantes na constituição desse modelo, que são os papéis dos participantes em uma conversa e o significado contido no discurso.

Assumindo que o discurso seja uma peça do comportamento linguístico que envolve múltiplos enunciados e participantes, que pode ser produzido por um ou mais participantes desse evento, pois em uma conversa mais de um dos envolvidos pode falar e compor diferentes enunciados do discurso, o modelo G&S não utiliza os termos *falante* e *ouvinte*, ou *locutor* e *interlocutor*, e sim os termos ICP – *initiating conversational participant* – e OCP – *other conversational participant* – para distinguir esses participantes. Em PB usaremos respectivamente participante iniciador da conversa (PI) e participante não iniciador da conversa (PNI).

Sobre o significado ou sentido do discurso é importante ressaltar que este é um problema ainda não resolvido, e que não é o foco desse modelo. O objetivo desse modelo é oferecer uma hipótese adequada sobre a *estrutura* do discurso, para que se possa desenvolver uma teoria adequada sobre o *significado* do discurso baseada nessa estrutura.

Sendo assim, podemos dizer que o discurso possui três componentes responsáveis por sua estruturação e interação: i) estrutura da sequência dos enunciados, ou estrutura linguística (*linguistic structure*); ii) a estrutura dos propósitos, ou estrutura das

intenções (*intentional structure*); e iii) o estado do foco de atenção, ou estado de atenção (*attentional state*). Juntos, esses três constituintes da estrutura do discurso suprem a informação necessária para que os participantes da conversa (PI e PNI) possam determinar como um enunciado se encaixa com outras partes do discurso, possibilitando que os participantes entendam por que algo foi dito e o que isso significa (Grosz & Sidner, 1986, p. 177), sem que se toque em aspectos sobre o significado do discurso como um todo.

2.1.1 Estrutura Linguística

A estrutura linguística, entendida no modelo G&S como a estrutura dos enunciados que compõem o discurso, é responsável por agregar tais enunciados em segmentos de discurso (doravante DS – *discourse segments*¹⁸).

O desafio desse modelo é a forma de delimitar os segmentos de discurso, pois embora diferentes pesquisas em diferentes teorias (Cohen & Levesque, 1980; Polanyi & Scha, 1986) tenham analisado uma variedade de gêneros de discurso não houve concordância suficiente sobre como delimitar as fronteiras dos segmentos de discurso. Existem várias pistas linguísticas que apontam para tais fronteiras, como pistas prosódicas tais como pausas, alongamentos (Chafe, 1979, 1980), taxa de elocução (Butterworth, 1975) e pistas textuais, como o uso de subordinações, coordenações, pronomes, referentes e marcadores discursivos (Hirshberg & Litman, 1993). Essas pistas textuais podem indicar inclusive as mudanças na estrutura das intenções ou o estado atencional do discurso. O exemplo de segmentação dos enunciados no quadro 1 a seguir, extraído do Corpus VoCE, é ilustrativo de como tais expressões linguísticas (em itálico colorido) funcionam como marcadores de fronteiras discursivas:

¹⁸ Serão usadas em inglês as siglas de alguns dos termos empregados nesta tese por uma facilidade de adequação com o material sobre o tema já publicado em inglês.

- (1) *mas hoje em dia* por exemplo eu tô lançando o u:: um meu disco novo no formato dum livro também
- (2) *então* vem o livro e o CD juntos
- (3) *que* é o Sambazz :::
- (4) *e o Samba/* e o livro explica a produção do disco
- (5) *então* além de ter o disco ainda tem o livro :
- (6) *que* eu escrevi todo ::: falando da produção do disco
- (7) *desd'* a:: das composições
- (8) *até* a finl/ finalização dele :::
- (7) *também* pra despertar um interesse na pessoa [assim] não simplesmente em baixar as o disco na in/ na internet de graça e :: ouvir as músicas ::
- (8) *mas* também pro cara : ter o:: ter a vontade de ler o livro :::
- (9) *e:: e* saber da história do disco *né?*

Quadro 1: Exemplo de expressões linguísticas como marcadores de fronteiras discursivas.

2.1.2 Estrutura das Intenções

A estrutura das intenções compreende os propósitos que subjazem ao discurso e seus componentes e suporta a distinção entre os propósitos fundamentais ao discurso.

Entre os participantes de uma conversa existe mais do que um único objetivo que os leva a participar de tal conversa. A distinção desses objetivos, ou intenções, é fundamental para o entendimento do discurso. Cada intenção que subjaz a um discurso em particular é chamada de propósito do discurso (doravante DP – *discourse purpose*).

De acordo com as proponentes do modelo, o propósito do discurso seria a intenção geradora do discurso. Essa intenção proporciona a razão do discurso ser produzido, assim como a razão do conteúdo particular do discurso ser sobreposta a qualquer outra informação (Grosz & Sidner, 1986). Sendo assim, cada segmento de discurso (doravante DS – *discourse segment*) carrega uma intenção, que é chamada propósito do segmento de discurso (doravante DSP – *discourse segment purpose*). O DSP especifica como cada segmento contribui para alcançar o propósito final do discurso. A figura 8 é ilustrativa dessas relações.

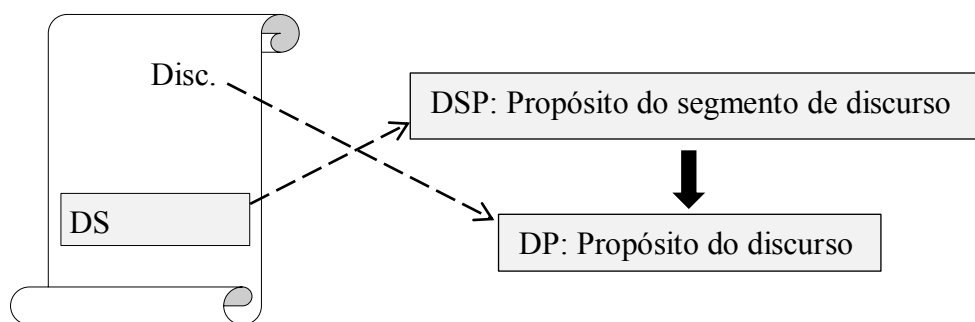


Figura 8: relação entre discurso e propósitos que subjazem ao discurso.

Na estrutura das intenções, um PI tem diferentes intenções que o levam a iniciar um discurso, no entanto essas intenções que servem como DPs ou DSPs se distinguem de outras intenções por terem o intuito de serem reconhecidas, enquanto outras intenções podem ser privadas, e não explicitadas nos DPs ou DSPs. O reconhecimento dos propósitos (DPs e DSPs) por parte dos participantes é essencial para alcançar o efeito pretendido pelo PI. Dentro dessa estrutura existem ainda duas relações estruturais que desempenham um papel importante na estrutura do discurso. Essas relações são denominadas pelo modelo G&S por *dominância e satisfação*.

Por exemplo, uma ação que satisfaz uma intenção em um determinado DSP(x) pode ser destinada a prover parte da intenção da ação em um DSP(y) precedente ou posterior. Nesse caso diríamos que DSP(x) *satisfaz* ou *contribui* com DSP(y), e que, portanto, DSP(y) *domina* DSP(x), o que gera uma hierarquia de dominância na estrutura dos DSs. A relação de satisfação entre os DSs pode ser significativa e também observável em casos em que o DSP(x) satisfaz o DSP(y). Porém, o DSP(x) tem que ser satisfeito antes de satisfazer o DSP(y) por um DSP(z), por exemplo. A relação de satisfação é um tipo de precedência que não exerce domínio sobre o DSP seguinte. O trecho segmentado em diferentes DSs no quadro 2 estabelece as relações entre seus DSPs descritas no quadro 3.

DS1 - (1) porque : tem uma con/ existia uma convicção de que o mundo ia acabar ::: né?
 (2) no dia que concid/ coincidisse isso :::

DS2 - (3) só que assim eh ::: a superstição o que que a superstição é?
 (4) ela nada mais é do que aquela velh' insegurança né?
 (5) então eu acredito em tudo :: eu todo mundo o que todo mundo me fala eu eu : coloco em ::

DS3 - (6) porque a a superstição ela vem :: de uma pessoa que num pensa ::: né?
 (7) se você for parar e pensar como é que um número :: pode ter poder e força em cima de você? :::

DS4 - (8) agora o que qu' ele tem poder e força?
 (9) porque na bíblia fica/ na bíblia sagrada católica apostólica romana fala:: que o número meia
 meia meia : é o número da besta

DS5 - (10) só que o número da best' ele num tem zero meia zero meia de dois mil e seis :
 (11) é o número meia meia meia ::
 (12) é diferente :: né? :::

DS6 - (13) então é por causa disso que acredito que há :: essas superstições

Quadro 2: segmentação hierárquica do discurso em diferentes DS.

DSP1 SATISFAZ DSP2	Existe uma superstição.
DSP2 DOMINA DSP3	A superstição é uma crença, um tipo de insegurança.
DSP2 DOMINA DSP4	
DSP3 SATISFAZ DSP4	Se você pensar vai entender que um número não tem poder sobre você.
DSP4 DOMINA DSP5	A bíblia cita o número 666 e não 6 do 6 de 2006.
DSP2 SATISFAZ DSP6	A superstição existe por esses motivos.
DSP3 SATISFAZ DSP6	
DSP4 SATISFAZ DSP6	
DSP5 SATISFAZ DSP6	

Quadro 3: relação de dominância e satisfação entre os DSPs do quadro 2.

O modelo G&S faz uma distinção importante entre a *determinação* e o *reconhecimento* de um DSP. A determinação se refere a uma relação semântica, que é a especificação completa do que é pretendido por alguém. O reconhecimento se refere a um processamento que permite a um participante do discurso identificar qual a intenção de outro (Grosz & Sidner, 1986). Esses conceitos são interrelacionados, pois a mesma

informação que determina um DSP pode ser usada por um participante para reconhecer a intenção deste DSP (Grosz & Sidner, 1986, p.179).

2.1.3 Estado de Atenção

E por fim, o estado de atenção é definido como uma propriedade intrínseca do discurso e não dos participantes do discurso. Esse estado é (i) inerentemente dinâmico, fazendo um registro de objetos, propriedades e relações salientes em cada ponto do discurso (Grosz & Sidner, 1986) e (ii) modelado por um conjunto de espaços focais (doravante FS – *focus space*), que são mudanças que ocorrem no estado de atenção motivadas por um conjunto de regras de transição que especificam as condições para se acrescentar ou excluir esses espaços. De acordo com o modelo G&S, a soma de todos os espaços focais em um discurso determina a estrutura focal, e o processo de manipulação desses espaços focais é chamado focalização. Os espaços focais incluem DSPs em seu escopo. Essa inclusão de propósitos no espaço focal mostra que os participantes da conversa focam não apenas o que está sendo dito, mas também porque isso está sendo dito (Grosz & Sidner, 1986).

O estado de atenção pode ser confundido muito facilmente com o estado cognitivo do participante, porém não é o caso. É importante esclarecer esta relação: o estado de atenção é apenas um dos componentes do que chamamos de estado cognitivo. O estado cognitivo compreende o estado interno do falante, que inclui conhecimento, crenças, desejos, emoções (raiva, medo, alegria), intenções, atitudes (paciência, impaciência, amizade) e atenção (*stress*, interesse, tédio). O estado cognitivo também inclui informações sobre o tema da conversa que pode ser expresso por meio do afeto (Scherer, 1984)

Dessa forma, o estado de atenção, por meio da estrutura focal, age de forma coordenada com a estrutura linguística e com a estrutura de intenções, ou seja, os DSs se ligam a seus respectivos FSs. Sendo assim, os primeiros DSs do discurso têm seus FSs menos acessíveis aos participantes, pois estes estão sendo sobrepostos pelo FSs dos últimos DSs a aparecerem no discurso por apresentarem propriedades, objetos e relações mais salientes aos participantes em comparação com os atributos dos primeiros DSs. A figura 9 a seguir exemplifica essa relação.

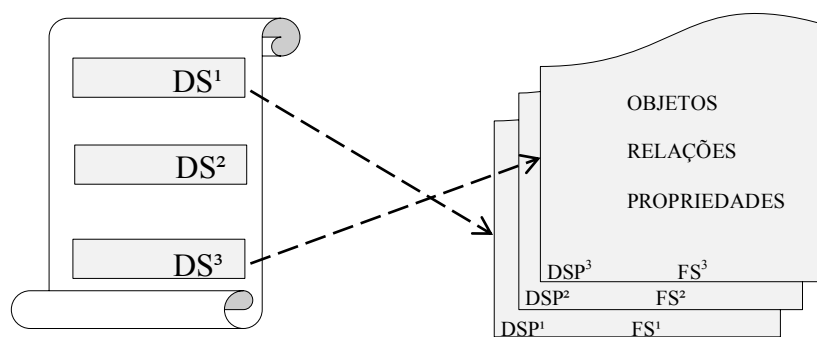


Figura 9: relação de saliência entre DSs e FSs.

A sobreposição dos espaços focais reflete a saliência relativa das entidades em cada espaço durante a porção correspondente ao segmento do discurso. Sendo assim, quando um DSP oferece uma nova intenção que se sobrepõe a outras na estrutura focal, espaços focais anteriores “descem” na pilha de sobreposições quando esse novo FS é inserido. Portanto, a estrutura focal – que compõe o estado de atenção – possui duas propriedades essenciais: i) ela parasita a estrutura de intenções, no sentido de que as relações entre os DSPs determinam a sobreposição dos FSs; e ii) assim como a estrutura intencional e a estrutura linguística, ela se desenvolve com o desenrolar do discurso, pois a estrutura das intenções se constrói pelos participantes com o progresso do discurso (Grosz & Sidner, 1986, p. 180).

Resumindo, a estrutura focal é o repositório para a informação contextual necessária para o processamento dos enunciados a cada parte do discurso. Ela distingue objetos, propriedades e relações que são salientes a cada momento e ainda oferece ligação para partes relevantes de ambas as estruturas linguística e de intenção. Durante o discurso uma crescente quantidade de informações é discutida, por isso é cada vez mais necessário ser possível identificar os segmentos relevantes para o discurso, as entidades que se fazem salientes e seus DSPs. O papel do estado de atenção em delinear a informação necessária para o entendimento é, portanto, central no processamento do discurso.

2.2 Aplicação do Modelo

O modelo G&S se aplica diretamente à análise da estrutura do discurso por meio da segmentação dos DSs e da construção de uma relação hierárquica entre esses segmentos.

O quadro 4 mostra a segmentação de um trecho de um discurso em 12 enunciados distintos, cujas fronteiras são determinadas pela presença de marcadores linguísticos (em *itálico colorido*). Cada um desses enunciados faz parte de um DS, que por sua vez possui um DSP, e este estabelece uma relação de satisfação ou dominância com outro DSP, por exemplo: DS2 e DS4 estabelecem relações de dominância com DS3 e DS5, respectivamente. No entanto DS3 estabelece relação de satisfação com DS4. A relação hierárquica entre os dados de fala segmentados no Corpus VoCE se encontra no Anexo II¹⁹.

- DS1** - (1) as outras visões ::: de criação do aparecimento das espécies ::: não são teorias científicas :::
(2) *dessa forma* é uma perda de tempo ficar comparando :::
(3) *o que* seria uma teoria científica ::: com uma teoria ::: não científica
- DS2** - (4) *eu* conheço ::: vários pesquisadores importantes :::
(5) *que* ::: tem as suas crenças religiosas :::
- DS3** - (6) *contudo* eles aceitam completamente a teoria evolutiva do Darwin :::
(7) *porque* não :: consideram ::: qu'exista incompatibilidade entre as duas :::
- DS4** - (8) *as teorias* são basicamente uma forma de ver o mundo ::
(9) *e* as teorias científicas são :: interessantes do pon'de vista de que ::: elas tem em período de duração ::
- DS5** - (10) *nesse aspecto* ::: a teoria do Darwin :: já tem cent'e cinquenta anos :::
(11) *isso* é um verdadeiro record ::
(12) *ele* praticamente só perde para teoria da gravidade de Newton

Quadro 4: relação hierárquica entre os DS que compõem o discurso.

A relação exposta na tabela 2 e na figura 10 são ilustrativas de como a teoria se aplica ao segmento discursivo apresentado no quadro 4. Cada um dos cinco DSs, formado

¹⁹ A segmentação do Corpus VoCE em segmentos discursivos foi feita por dois pesquisadores (a autora da tese e uma mestranda em linguística) levando em consideração apenas a transcrição das falas. Ao final das transcrições estas foram comparadas e chegou-se a uma versão final.

por um conjunto de enunciados, tem a intenção (DSP), por parte do iniciador da conversa, de que os outros participantes acreditem ou aceitem suas proposições.

DS	Enunciados	Intenções
DS1	as outras visões ::: de criação do aparecimento das espécies ::: não são teorias científicas ::: <i>dessa forma</i> é uma perda de tempo ficar comparando ::: <i>o que</i> seria uma teoria científica :::: com uma teoria :::: não científica	existe somente uma teoria científica válida para o aparecimento das espécies;
DS2	<i>eu</i> conheço ::: vários pesquisadores importantes ::: <i>que</i> ::: tem as suas crenças religiosas :::	existem pesquisadores com crenças religiosas;
DS3	<i>contudo</i> eles aceitam completamente a teoria evolutiva do Darwin ::: <i>porque</i> não :: consideram ::: qu'exista incompatibilidade entre as duas	as crenças religiosas não são incompatíveis em aceitar a teoria do Darwin como válida;
DS4	<i>as teorias</i> são basicamente uma forma de ver o mundo :: <i>e</i> as teorias científicas são :: interessantes do pon'de vista de que ::: elas tem em período de duração ::	teorias têm períodos de duração (e podem ser falseadas);
DS5	<i>nesse aspecto</i> ::: a teoria do Darwin :: já tem cent'e cinquenta anos ::: <i>isso</i> é um verdadeiro record :: <i>ele</i> praticamente só perde para teoria da gravidade de Newton	a teoria darwiniana perdura por 150 anos, é uma teoria forte e não foi falseada até hoje

Tabela 2: composição dos DSs e seus DSPs.

Seguindo o modelo G&S, o espaço focal do DSP5 assume a posição mais alta na posição dos FSs no momento em que ocorre no discurso, enquanto o espaço focal de DSP1 fica sobreposto por todos os outros (ver figura 10). Nesse exemplo, ele inicia o discurso com foco sobre a validade da teoria do surgimento das espécies, e termina com foco direcionado para a robustez da teoria, ou seja, a intenção do discurso até atingir o FS5 vai sendo construída pela sobreposição dos espaços focais na atenção do PNI, e pela relação de dominância de cada uma das intenções dos DSPs.

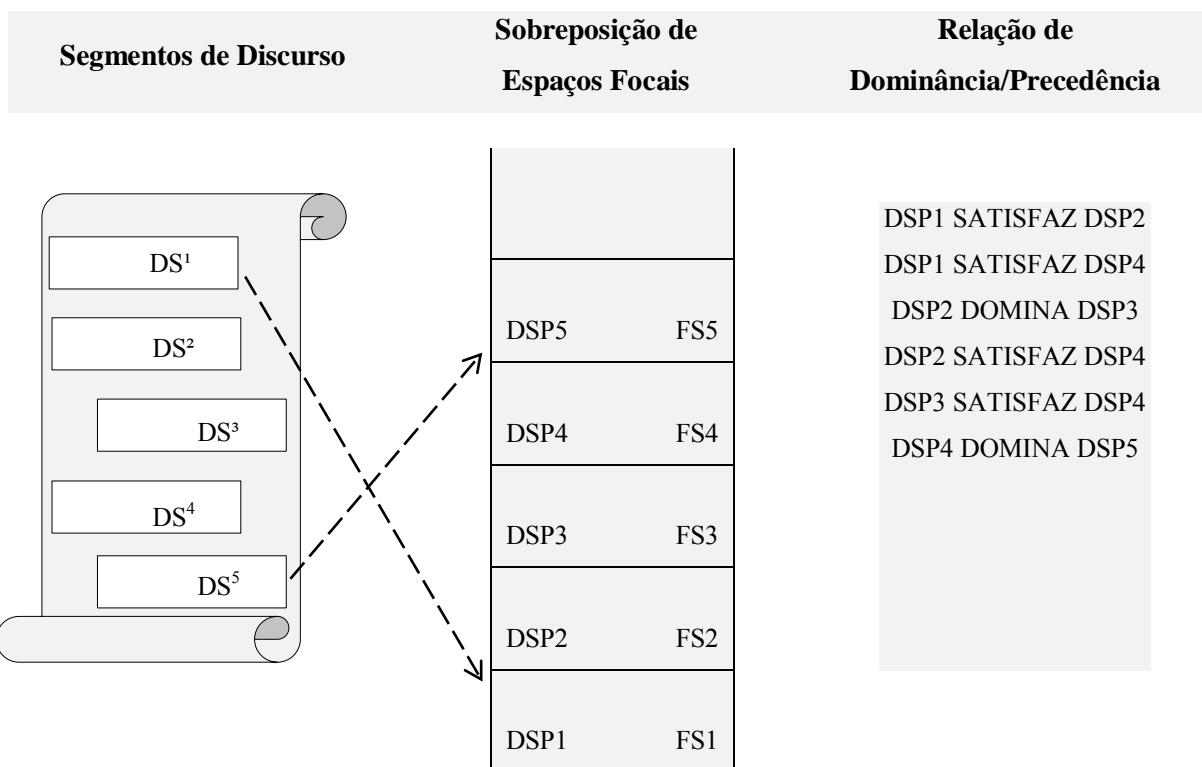


Figura 10: sobreposição dos FS e relação de dominância e satisfação entre DSs.

2.3 Elementos de Construção do Discurso

O modelo G&S, em sua aplicação, se utiliza de elementos de construção do discurso que são empregados na segmentação dos enunciados que compõem os DPs e também no reconhecimento das intenções dos participantes. Os elementos mais explorados em Grosz & Sidner (1986) são as interrupções no discurso e o uso de marcadores linguísticos, como marcadores discursivos, pronomes, coordenações, subordinações, referentes, pausas e hesitações. Na abordagem que se segue, esses elementos de construção do discurso são ampliadas segundo a perspectiva textual-interativa da conversação. De acordo com a gramática de Jubran e Koch (2006), as interrupções, de modo geral, contêm,

ou expressam, elementos como as hesitações, repetições, correções, parafraseamentos e parentetizações. Paralelamente aos elementos relacionados à interrupção, há elementos independentes de interrupções como a tematização/rematização, a referenciação e o uso de marcadores discursivos, a última também explorada em Grosz & Sidner (1986).

2.3.1 Interrupções

A interrupção é um fenômeno intrínseco da oralidade, caracterizado por um inacabamento ou qualquer tipo de ruptura no desenvolvimento do enunciado, que pode ser retomado com ou sem modificações na estrutura do discurso (Silva & Crescitelli, 2006).

Quando não ocasionadas por fatores externos, como por exemplo, quando um participante da conversa interrompe outro, as rupturas evidenciam uma falha no processamento do discurso, podendo ocorrer nos níveis sintático e lexical. Se por um lado a interrupção reflete uma ruptura na fluidez do discurso, sua presença é marcada também pela retomada do discurso. A retomada funciona como um retorno semântico a um DSP anteriormente interrompido.

A interrupção seguida por uma retomada tem a função de sinalizar prospectivamente a ocorrência imediata a ela de diferentes estratégias de construção do discurso, como, por exemplo, a correção, a repetição, os parênteses e a paráfrase. (Silva & Crescitelli, 2006)

Grosz & Sidner (1986) propõem duas definições *ligeiramente diferentes* (como as próprias autoras definem) para as interrupções, independentemente de quais tipos de interrupções se tratem. Tais interrupções são classificadas como forte ou fraca. A interrupção é chamada forte quando ocorre em um DS cujo DSP não é dominado ou satisfeito por qualquer segmento precedente. A interrupção fraca ocorre quando o DS interrompido apresenta um DSP que não é dominado nem satisfeito por um segmento *imediatamente* precedente. Esta definição do modelo G&S não prevê a retomada do discurso, pois a retomada acaba por eliminar em alguns casos a ausência das relações de dominância e satisfação entre os DSPs. Os exemplos a seguir mostram dois casos de interrupção com exemplos extraídos do Corpus VoCE:

i)

1- (1) *pausa* só pra todo mundo entender

2 - 3 - (2) você vai baixar um vídeo pra você assistir ::

(3) aí você baixa ele inteiro pela internet

(4) depois você bota um play lá : né [isso] pra tocar e passa

ii)

1- (1) olha : *é o seguinte quando/ eu me lembro perfeitamente*

(2) aliás esse ano fo foi :: ah comemorou-se quinze anos de São Paulo Fashion Week ::

(3) portanto né? faz faz ::: eh muito tempo [risos] ehhhh :: ehhhh melhorou muito ::

2 - (4) as empresas que se apresentavam primeiro copiavam muita roupa fora ::

As estratégias de construção do discurso que se utilizam de interrupções no desenvolvimento de enunciados serão apresentadas a seguir.

2.3.1.1 Hesitação

A hesitação, considerada aqui com um tipo de interrupção, é um fenômeno intrínseco da oralidade, e sobretudo, como afirma Marcuschi (2006), é um fenômeno de processamento da fala, pois revela os recursos usados pelos falantes para resolverem problemas que surgem devido ao processamento *on-line* das formas e conteúdos do discurso. O modelo G&S, assim como a perspectiva textual-interativa, pressupõe o processamento do discurso no momento em que este está sendo produzido.

A hesitação envolve aspectos formais, que se dividem entre: fenômenos prosódicos, como pausas (silêncio intraturno ou *switching pause*); alongamentos vocálicos (típico da construção de listas); expressões hesitativas (*ah, éh, mm*); itens funcionais, como artigos, preposições, conjunções e pronomes (*pra... pra; de...de...; um...*); itens lexicais, geralmente verbos (*há.. há..*); marcadores discursivos (*bom...; sei lá...; sabe...; né...*); e fragmentos lexicais, ocorrendo em geral com item lexical duvidoso ou de difícil acesso no momento de formulação do discurso (*a/ a/ acabou; con/ ah... construindo*).

Ainda sobre o fenômeno da hesitação, Merlo (2012) vê estas como dinâmicas e inerentes ao discurso, atuando de forma a torná-lo mais coerente, pois as hesitações

indicam possibilidades de escolhas linguísticas e diversidade lexical, e acabam por evitar a cristalização dos enunciados (Merlo, 2012, p.18).

Os cinco exemplos a seguir mostram a ocorrência de hesitações e sua relação com os DSs, que estão indicados por identações distintas.

iii)

- 1- (1) olha : é o seguinte quando/ eu me lembro perfeitamente
- (2) aliás esse ano **foi foi :: ah** comemorou-se quinze anos de São Paulo Fashion Week ::
- (3) portanto né? **faz faz ::: eh** muito tempo [risos] **ehh :: ehh** melhorou muito :

iv)

- 2- (4) não precisa dizer mais nada ::
- (5) acho que aquilo ali se **eh uh** as pessoas falam por si só :::
- 3 - (6)eu acho que a gente tem que ser imparcial :::
- 4 - (7) **eh::: oh::** jornalismo tem isso né? a gente tem que dar : todos os lados né?

v)

- 1 - (1) **eh eh** uma coisa *de* ::: **ehh :: de ser ::: eh** : informal :: de ser leve de ser alegre de ser uma coisa :::
- (2) e não é associado a luxo

vi)

- 2 - (9) então assim eu sinceramente acho que não **cabe :::** a nós : âncoras
- (10) acho que cabe ao público julgar aquilo
- 4 - (11) até porque :: eu tenho uma :: formação diferente eu tenho uma cultura diferente cada um tem **né? :: a a o os** seus princípios :::

vii)

- 1 - (1)**eh eh eu f/ eu fico** eu passo ali perto da USP quase todo dia né? :
- 2 - (2) **eh :::** os caras ali aconteceu uma coisa interessante aí do estado brasileiro né? :
- (3) **pra/ eh p/ por** conta **do do** Serra também :: a gente tem um governador novo né?
- 3 - (4) ninguém sabe quem ele é :::
- (5) mas eh basicamente o que ele tem feito hoje em dia é não fazer nada :::

2.3.1.2 Repetição

Funcionalmente, a definição de repetição, como uma forma de interrupção é a produção de segmentos textuais idênticos ou semelhantes, duas ou mais vezes no âmbito de

um mesmo evento comunicativo e que interrompem o fluxo do discurso (Marcuschi, 2006). A repetição contribui para a organização discursiva, ocorrendo muitas vezes associada a hesitações, pois assim como esta, a repetição é um fenômeno decorrente do processamento da fala no momento da produção na medida em que atua na monitoração da estrutura linguística e intencional do discurso.

Iremos abordar aqui as formas de repetição ligadas à interrupção e que se relacionam à estrutura intencional do discurso, responsável pela delimitação dos DSP. Dessa forma, os aspectos da repetição relacionados à ênfase e aspectos sociais do discurso não serão abordadas.

As repetições exemplificadas aqui são, respectivamente, do tipo: lexical, fonológica, e oracional.

viii) *lexical*

- 1 - (3) o auê com a moda : é imenso ::: eh :: **vende** muito programa de rádio de televisão **vende** pro/ **vende** revista ::: eh
- 2 - (4) é um produto de entertainment da pesada moda
- (5) porque o povo :: ama : ver : desfile : **é bonito** mesmo desfile de moda : **é bonito** de ver ::
É engraçado é interessante tudo:::

ix) *fonológica*

- 3 - (6) que eu escrevi todo ::: falando da produção do disco desd' a:: das composições até *a finl/ finalização* dele
- (7) também pra despertar um interesse na pessoa [assim] não simplesmente em baixar as o disco **na in/ na internet** de graça e :: ouvir as músicas ::
- (8) mas também pro cara : **ter o:: ter a** vontade de ler o livro ::: e:: e saber da história do disco né?

x) *oracional*

- 1 - (1) **acho que sim ::**
- (2) a gente consegue perceber isso pelos e-mails que a gente recebe :: ehh pelas ligações pro CAT
- 2 - (3) que é Central de Atendimento ao Telespectador :: da TV Globo
- (4) **eu acho que sim :::**
- (5) **acho que acho que** cada vez mais :: o nosso público

2.3.1.3 Correção

A correção, assim como as outras estratégias de construção do discurso discutidas aqui, é reveladora do processamento do discurso e do monitoramento deste. A função da correção, portanto, é reformular partes do discurso em que há evidência de falhas na formulação. A correção representa uma atividade intencional, que é a de reestruturar os DSs em um discurso.

As correções podem ser confundidas ou fundidas com outras estratégias como a hesitação, por exemplo. A diferença fundamental neste caso é que a hesitação ocorre quando o “problema” de formulação é detectado no momento da produção, ou seja, *on line*, tendo assim um aspecto prospectivo (Fávero et al, 2006). Enquanto na correção ocorre uma reformulação de partes de um DS por considerá-las inadequadas no processamento da fala (Fávero et al, 2006), e, portanto, invalidando a intenção desse DS ou retificando-a por outra intenção no discurso.

As correções podem ser de ordem: fonético-fonológica, quando se corrige pronúncia ou articulação; lexical, quando a seleção lexical não era a pretendida ou a mais adequada; morfossintática, quando ocorrem falhas de concordância, regência, etc. Os exemplos do Corpus VoCE ilustram esses tipos de correções:

xi) fonético-fonológica:

2 - (2) uma coisa que eu discordo *veement/veementemente* é eu não acho que eu tenho vergonha de ser brasileiro como muitos dos ouvintes começaram colocar

xii) fonético-fonológica:

3 - (6) eu fui ver por exemplo ah essa essa que ce tá falando essa essa coleção
4 - (7) que era da Reserva *uma masc/ marca masculina* :: Carioca ::
5 - (8) e eu também fiquei olhando pro show do Lobão

xiii) lexical:

- 1- (1) porque : **tem uma con/ existia uma convicção** de que o mundo ia acabar ::: né?
- (2) no dia que concid/ coincidissem isso :::

xiv) lexical:

- 1 - (1) mais hoje em dia por exemplo eu to lançando o u::: um meu disco novo no formato dum livro também
- (2) então vem o livro e o CD juntos
- (3) que é o Sambazz :::
- 2 - (4) **e o Samba/ e o livro** explica a produção do disco

xv) morfossintática:

- 5 - (7) então as réplicas elas são constantes ::
- (8) ce não pode ah ah adequar que a sua réplica um dia **foi um ::: uma inocência** né? :::
- (9) então ::: os treplicantes são aquelas crianças que vão defender o futuro Aton

2.3.1.4 Parafraseamento

A paráfrase, assim como a correção, é uma reformulação do discurso, e tem como intuito reelaborar uma mensagem para que esta seja mais explícita, clara ou mesmo mais enfática. A distinção entre paráfrase e correção (apesar de ambas apresentarem reformulações e aspecto retrospectivo) se baseia no fato de que a paráfrase apresenta uma reformulação do discurso com o intuito de reelaborar uma mensagem para que esta seja mais explícita, clara ou mesmo mais enfática. Com a correção ocorre a reformulação de partes de um DSP considerado inadequado pelo falante no processamento da fala (Fávero et al, 2006).

A paráfrase pode ser anunciada por marcadores discursivos que indicam o parentesco semântico entre o DS que contém paráfrase e o DS de origem, ao qual se liga. Esses marcadores podem ser específicos, como *em outras palavras*, *como disse a pouco*, *resumindo*, ou mais abrangentes, pois não constituem um índice exclusivo de ligação de equivalência entre os enunciados, como: *então*, *quer dizer*, *ou seja*, etc. Seguem exemplos do Corpus VoCE para ilustrar o parafrasamento.

xvi)

- 1- (2) ah::: é baseada nnn em algo que o Apple TV já vinha fazendo
(3) que eh :: o consumo streaming de vídeo
2- (4) ***quer dizer você vai assistir o vídeo vindo sob demanda da internet no momento
:: em que você está assistindo***

xvii)

- 4 - (5) o streaming que ela tá falando é você vê ao mesmo tempo :::
(6) ***ce vê dois minutos ::: depois dele ser baixado***
(7) entendeu?

xviii)

- 1 - (1) as pessoas acham que com esse com esse aumento de consumo que teve especialmente da classe C D ::
(2) ***que a indústria brasileira se beneficiou ::***
2 - (3) ***e vou dizer que infelizmente não foi isso que aconteceu ::***
3 - (4) o comércio sim :: e o consumidor sim tiveram mais oh mais : opções [aham]
(5) mas a indústria brasileira não
4 - (6) por quê? porque nós não temos preço
5 - (7) então essa roupa toda que veio pra abastecer :: com preço bom ::
(8) pra pra abastecer classe C e D :: veio da China e da Índia :::
6 - (9) ***ou seja não foi a indústria brasileira que se beneficiou disso***

xix)

- 3 - (15) então ***quando eu comentei*** :: que nem o dia/ o demônio aguenta ::
(16) ***quer dizer :: nem o demônio aguenta deus : né***

2.3.1.5 Parentetização

Os parênteses são uma forma de interrupção do discurso em que se inserem breves desvios, ou informações paralelas, em um determinado tópico discursivo e que não afetam a coesão do segmento dentro do qual se inserem (Jubran, 2006). Os parênteses não só não afetam a coesão do DSP em que se inserem como promovem avaliações e comentários laterais sobre o que está sendo dito e sobre como se diz (Jubran, 2006).

Esse tipo de inserção que os parênteses operam no discurso possui uma função pragmática, pois se apresentam como recursos pelos quais o participante insere informações materiais que interrompem o discurso avaliando e informando a posição dos participantes nesse discurso (Jubran, 2006).

Os parênteses possuem marcas formais de inserção, que podem ser observadas no segmento parentético pela ausência de conectores lógicos, e marcas prosódicas, como mudanças na taxa de elocução (cf. Tenani, 1996). No contexto do discurso essas marcas podem ser de interrupção (pausas não preenchidas, suspensões) e de reintrodução (pausas não preenchidas, uso de conectores, repetição de itens lexicais) (cf. Jubran, 2006, p.310).

Nos exemplos a seguir podemos ver como os parênteses constituem novos DSPs, como transmitem as intenções avaliadoras e informativas particulares dos PI sobre o DP, e as marcas de inserção pela ausência de conectores lógicos.

xx)

0 -1 - (1) olha : é o seguinte quando/ eu me lembro perfeitamente

(2) *aliás esse ano fo foi :: ah comemorou-se quinze anos de São Paulo Fashion Week::*

(3) *portanto né? faz faz ::: eh muito tempo [risos] ehhhh :: ehhhh melhorou muito ::*

2 - (4) as empresas que se apresentavam primeiro copiavam muita roupa fora ::

(5) a gente via no confronto na passarela :: uma duas três vezes a mesma roupa copiada fora

xxi)

2 - 3 - (3) veja bem nós temos uma imagem

(4) *sobretudo na moda né? :::*

4 - (5) eles admitem eles :: eles : olham com prazer a nossa roupa esportiva :: biquini eh:

xxii)

3 - (4) então por isso que muitas pessoas tem muitos me/ muito medo das coisas e as vezes não sabe onde vem :::

(5) e as vezes vem dessas pequenas coisas ::

4 - (6) *eu não sei você mas eu tive um tempo na minha vida quando era mais nova a gente tinha muito contato com dono de fazenda fazendeiro essas coisa(s) ::*

(7) *e eles contavam aquelas coisas né dos escravos que arrastava corrente na madrugada*

5 - (8) *então você ia p'essas fazendas pra dormi quem disse que ce dormia? :::*

(9) *se passava a noite inteira com medo que se ia ouvi corrente*

(10) aí você acabava ouvindo

(11) porque o seu medo faz você ouvir :::: né o medo faz a gent/ :

xxiii)

4 - (9) agora o Brasil tem uma imagem tão positiva lá fora ::

(10) *você pode fazer a experiência*

(11) *eu ja fiz essa experiência :*

(12) *todo mundo já fez :::*

5 - (13) *você chega lá fora oh ou você tá diante dum estrangeiro*

(14) *você diz que ce é brasileira ::*

(15) *todo mundo sorri :::*

(16) *imediatamente as pessoas sorriem [é verdade]*

(17) *vai dizer que ce é Belga ohh ohh qualquer coisa ninguém ri*

6 - (18) então : a imagem do Brasil : é uma imagem solar agradável simpática e tudo ::

(19) nós temos que aproveitar ess'imagem

2.3.2 Elementos de construção do discurso sem interrupção

Da série de elementos usados na construção do discurso nem todos se classificam como interrupções em seu desenvolvimento temporal. Alguns elementos importantes e também abordados pela perspectiva textual interativa (Marcuschi e Koch, 2006) envolvem o *status* da informação, ou tematização e rematização, a referenciação e o uso de marcadores discursivos.

2.3.2.1. *Status* da informação

Um dos fatores que suportam a estrutura do discurso está em como identificamos o *status* da informação no discurso. Por *status* da informação entendemos

aqui a distinção da informação entre nova e dada. A informação chamada *nova* introduz um novo elemento no discurso, e, portanto, possui formas de marcar e chamar atenção para essa inserção. O mesmo acontece quando essa informação é reintroduzida no discurso, como informação *dada*, pois também necessita de estratégias que marquem seu *status*.

Alguns autores relacionam o *status* da informação com a posição em que a informação se encontra na divisão tema-rema dos enunciados. De acordo com essa divisão, cada enunciado se divide em dois segmentos, um (tema, estático) para o qual é atribuída a predicação trazida por outro segmento (rema, dinâmico). Portanto, a informação nova assume a rematização e a informação dada à tematização nos enunciados (Moraes, 1998, Koch, 2006).

Diversos autores trabalharam com o status da informação no discurso defendendo que a marcação deste não se deve apenas a fatores entoacionais (Prince, 1981; Chafe, 1976; Brown, 1983; Baumann, 2006; Baumann & Riester, 2010). Os trabalhos de Gravano & Hirschberg (2006), Fernandes (2007) e Lucente, et al (2012) mostram que além da entoação a posição das informações novas e dada na estrutura discursiva também relevantes. Em Lucente et al (2012) esse fato fica evidente no PB após a realização de uma série de experimentos com informação nova e dada na fala espontânea. Os resultados desse trabalho mostraram que no PB (i) os falantes usam a função gramatical e a estrutura do discurso para marcar o *status* da informação, e que a (ii) informação dada nem sempre aparece desacentuada no PB, assim como a informação nova não aparece proeminente em todas as situações, como preveem os trabalhos mais clássicos sobre o assunto (Prince, 1981; Clark & Haviland, 1977; Brown, 1983).

Os exemplos a seguir mostram a presença de informações novas (**negrito**) e dadas (**negrito**, *itálico*) no discurso espontâneo.

xxiv)

- 1 - (1) eh eh uma coisa de ::: ehh :: de ser ::: eh : informal :: de ser leve de ser alegre de ser uma coisa/
(2) e não é associado a luxo
2 - 3 - (3) veja bem nós temos uma **imagem**
(4) sobretudo na moda né? :::
6 - (5) eles admitem eles :: eles : olham com prazer a nossa roupa esportiva ::
biquini eh :
(...)
4 - (9) agora o Brasil tem uma **imagem** tão positiva lá fora ::
(10) você pode fazer a experiência
(11) eu já fiz essa experiência :
(...)
5 - (18) então : a **imagem** do Brasil : é uma imagem solar agradável simpática e tudo ::
(19) nós temos que aproveitar essa **imagem**
(20) e colocar esses elementos na nossa moda
(21) pra qu'ela seja vista da mesma maneira

xxv)

- 1 - (1) as pessoas acham que com esse com esse aumento de consumo que teve especialmente da classe C D ::
(2) que a **indústria brasileira** se beneficiou ::
2 - (3) e vou dizer que infelizmente não foi isso que aconteceu ::
3 - (4) o comércio sim :: e o consumidor sim tiveram mais oh mais : opções [aham]
(5) mas a **indústria brasileira** não,
4 - (6) por quê? porque nós não temos preço
5 - (7) então essa roupa toda que veio pra abastecer :: com preço bom ::
(8) pra pra abastecer classe C e D :: veio da China e da Índia :::
6 - (9) ou seja não foi a **indústria brasileira** que se beneficiou disso

xxvi)

- 1 - (1) eu acho que ainda existe um pouco de confusão [uhum] da população como toda lei ao ser ::: ah entrar em vigor :::
(2) porque :: os casos da lei Maria da Penha são relacionados à **violência doméstica** :::
2 - (3) então pras pessoas entenderem a caracterização do que é uma **violência doméstica** ::: eles acabam confundindo
(4) porque a lei :: ela dá margem ::: algumas interpretações
(5) mas ao mesmo tempo ela está sendo mais a/ ::: aplicada pra lei ::: pra **violência doméstica**

xxvii)

- 1 - (1) eh eh eu f/ eu fico eu passo ali perto da **USP** quase todo dia né? :
- 2 - (2) eh ::: os caras ali aconteceu uma coisa interessante ai do estado brasileiro né?
- (3) pra/ eh p/ por conta do do Serra também :: a gente tem um governador novo né?
- 3 - (4) ninguém sabe quem ele é :::
- (5) mas eh basicamente o que ele tem feito hoje em dia é não fazer nada :::
- 4 - (6) e a gent' também tem um reitor na **USP** que é mais ou menos o mesmo mesmo pacote ::
- 5 - (7) ai aconteceu uma coisa interessante essa semana que foi a privatização do portão da **USP** né? :::
- (8) porque né? a **USP** tem vários portões tem o portão principal :::

2.3.2.2 Referenciação

A referenciação é um dos elementos que não pressupõe a interrupção do fluxo discursivo, porém, assim como a correção e a parentetização, ela diz respeito às operações efetuadas pelos participantes à medida que o discurso se desenvolve (Marcuschi, 2006). Por referenciação entende-se o objetivo de referir, remeter ou retomar (cf. Marcuschi e Koch, 2006, p.383) algo já presente no discurso, portanto, é um elemento fundamental para a organização e hierarquização dos DSPs.

A noção de referenciação adotada por Marcuschi e Koch (2006) leva em consideração: (i) estratégias de referenciação, como a rotulação que “promove uma recategorização da informação precedente por meio de novas predicções atributivas, ajustando o saber disponível a respeito do objeto-de-discurso” (Marcuschi e Koch, 2006, p.392), e a argumentação, que imprime uma orientação argumentativa ao recategorizar o referente; (ii) estratégias de progressão referencial, por meio de descrição definida, nominalização, pronominalização e associação; e (iii) processos de referenciação anafórica. Os exemplos a seguir ilustram algumas dessas estratégias de referenciação.

xxviii) *rotulação*:

- 1- (1) mas não tem : uma Gisele Bundchen tem [] [] **a Gisele Bundchen**
- (2) num tem duas nem três tem **a Gisele!**
- 2- (3) então quando você chama **uma : top desse jeito** :: de fato aquilo levanta bastante o de/
o desfile

xxix) *argumentação*:

- 1 - (1) as outras visões ::: de criação do aparecimento das espécies ::: não são teorias científicas :::
- (2) **nessa forma** é uma perda de tempo ficar comparando ::: o que seria uma teoria científica :::
- com uma teoria :::: não científica
- 2 - (...)
- (7) e as teorias científicas são :: interessantes do ponto de vista de que ::: elas tem em período de
duração ::
- 4 - (8) **nesse aspecto** ::: a teoria do Darwin :: já tem cent'e cinquenta anos :::

xxx) *pronominalização*:

- 1- (1) eu comecei a estudar teatro com dezessete anos :::
- (2) e na casa das sete mulheres eu já tava com vinte sete :::
- 3- (3) e **nesse período** : eh eu estudei continuo estudando até hoje

2.3.2.3 Marcadores Discursivos

Assim como no modelo G&S, em Hirschberg & Litman (1993)²⁰ os marcadores discursivos são interpretados como marcadores funcionais diretos da estrutura do discurso, pois aparecem marcando a introdução de um novo tópico ou subtópico, retomadas, parafraseamento, parentetização, digressões, ou seja, possuem a função de orientadores da interação (Risso et al, 2006) e desempenham papel importante na organização discursiva.

Essenciais para a estrutura discursiva, os marcadores são definidos, de acordo com a perspectiva textual-interativa, como sendo elementos de constituição diversificada

²⁰ Hirschberg & Litman (1993) usam o modelo G&S para a segmentação dos dados e análise de marcadores discursivos.

que envolvem no plano verbal sons não lexicalizados, palavras, locuções e sintagmas, constituindo-se como uma categoria pragmática no funcionamento da linguagem (Risso et al, 2006).

Os exemplos a seguir mostram a ocorrência de marcadores discursivos isolados no Corpus VoCE.

xxxì)

olha : é o seguinte quando/ eu me lembro perfeitamente

olha eu queria chorar porque eu não tenho **assim ahh** condição ::: de prever o meu futuro :: além do ponto final se tornando uma outra língua brasileira :::

então na verdade um reator nuclear é uma termoelétrica digamos metida a besta né?

então ::: **eh** a Trama junto com outras gravadoras independentes começaram apostar em outras coisas

então eh duas coisas em relação aos ouvintes que mandaram qu/ é milhões de críticas em respeito ao ao Renan

cara na verdade assim a a Trama ela ela ::: deu muito certo porque além d/ de ser uma gravador' independente ::: eh co' a mentalidade gravador' independente

ehh na verdade assim não é um cara que dá opinião **né?**

quer dizer você vai assistir o vídeo vindo sob demanda da internet no momento :: em que você está assistindo

quer dizer que a gente fala é que :: **ehh** o auê é muito maior do que ah :: ah : ah do que o faturamento vamos dizer **assim né?** :::

quer dizer :: nem o demônio aguenta deus : **né?**

agora quando você chama um Lobão :: você distrai o olhar

agora o Brasil tem uma imagem tão positiva lá fora ::

se tá na sua pele ele vai acabar difundindo **né?** :::

até porque :: eu tenho uma :: formação diferente eu tenho uma cultura diferente cada um tem **né?** :: a a o os seus princípios :::

bom explodiu aliviou a pressão foi embor' o calor acabou **né?** :::

2.4 Discussão

Os elementos que compõem a estrutura do discurso no modelo G&S podem agir interrompendo o fluxo do discurso ou não. De acordo com esse ponto de vista, no discurso ou na fala, as ideias se estruturam em unidades hierárquicas, não necessariamente longas ou completas do ponto de vista sintático, mas perfeitamente analisáveis, pois possuem elementos construtores que agem possibilitando o entendimento e a análise.

Tais elementos são parte integrante e fundamental dos segmentos discursivos, eles marcam pontualmente trechos do discurso e têm sua função comunicativa clara para as partes envolvidas em uma conversa. No modelo G&S, os elementos de construção além de funcionarem como indicadores dos locais em que os DSPs se segmentam e se organizam hierarquicamente seguindo um critério semântico, também apresentam propriedades prosódicas que influenciam a relação entre a estrutura discursiva e a entoacional, como se verá adiante.

2.5 Sumário

- o modelo G&S pressupõe três componentes responsáveis pela estruturação e interação do discurso: estrutura linguística; estrutura das intenções; e estado de atenção;
- a estrutura linguística agrega os enunciados em segmentos do discurso (DS);
- a estrutura das intenções compreende os propósitos que subjazem ao discurso, que são o propósito do discurso (DP) e os propósitos dos segmentos do discurso (DSP);
- o estado de atenção faz um registro dos objetos, propriedades e relações salientes em cada parte do discurso, determinados por espaços focais (FS);
- os DSPs estabelecem relações de precedência e dominância entre si; essas relações se organizam a fim de que os participantes de uma conversa entendam e aceitem as proposições do discurso;

- a estrutura do discurso conta com elementos que podem se realizar por meio de interrupções – hesitação, repetição, correção, parafraseamento e parentetização – ou sem interrupções no fluxo discursivo – status da informação, referenciação e marcadores discursivos;
- os elementos de construção do discurso auxiliam na segmentação dos DSPs em unidades hierárquicas perfeitamente analisáveis ancoradas nos aspectos: atenção, intenção e estrutura.

III ASPECTOS DINÂMICOS DA FALA E DA ENTOAÇÃO

1. NOTAÇÃO E DESCRIÇÃO ENTOACIONAL

Para a notação e descrição dos padrões entoacionais do Corpus VoCE foi utilizado o sistema DaTo (acrônimo de *Dynamic Tones of Brazilian Portuguese Intonation*) de notação entoacional, proposto por Lucente (2008). O sistema DaTo propõe para notação entoacional um conjunto de contornos dinâmicos e de níveis de fronteira que compõem o quadro entoacional do PB. Paralelamente aos fenômenos entoacionais, o sistema DaTo apresenta segmentação manual da fala em segmentos funcionais e a segmentação semiautomática dos enunciados em unidades rítmicas.

O sistema DaTo, por ter como foco uma notação entoacional baseada na avaliação auditiva de proeminências entoacionais, na observação visual da curva de f0 e na notação manual, remete inevitavelmente à teoria métrica-autossegmental (Lieberman, 1975; Pierrehumbert, 1980), ou teoria MA, que tem como expoente o sistema ToBI (*Tones and Break Indices*) de notação entoacional, baseado no trabalho de Pierrehumbert (1980), desenvolvido inicialmente para a notação do inglês americano (Beckman et al, 1994; Silverman et al, 1992)

Ao contrário do que propõe a teoria MA, a notação DaTo segue uma perspectiva dinâmica (Xu, 1999; 2005; 2006; Barbosa, 2006) ao analisar os fenômenos da entoação. Essa perspectiva dinâmica contrapõe-se a uma notação métrico-linear, como no sistema ToBI, na medida em que procura não representar a estrutura tonal por meio de eventos tonais isolados, mas por meio de contornos entoacionais dinâmicos, que se relacionam ao longo da estrutura, e são determinados por aspectos fonéticos como duração, intensidade (o que Xu (2005) chama de *strength*), gama de variação tonal e alinhamento com o material linguístico (Lucente, 2008a).

1.1 Teoria MA

De acordo com Ladd (1996a), a teoria MA propõe-se a caracterizar adequadamente na estrutura tonal, elementos categoricamente distintos e promover o mapeamento dos parâmetros acústicos contínuos a partir de elementos fonológicos, ou seja, a partir de representações fonológicas.

Os fundamentos da teoria MA são: i) a linearidade da estrutura tonal; ii) a distinção entre proeminência (*pitch accent*) e acento lexical (*stress*); iii) a análise dos *pitch accents*²¹ em relação à altura dos tons; iv) e a relação entre características locais e tendências gerais.

Partindo da hipótese da linearidade da estrutura tonal, a teoria MA representa fonologicamente os contornos entoacionais como sequências de eventos entoacionais discretos (Ladd, 1996a, p. 43). Sendo assim, a teoria assume claramente a distinção entre eventos e transições que ocorrem na estrutura tonal, reconhecendo partes da estrutura como linguisticamente importantes e outras como apenas transições preenchedoras entre eventos localizados. Os eventos considerados importantes do ponto de vista linguístico são os *pitch accents*.

Para essa teoria, as relações de proeminência são em primeira instância abstrações fonológicas, que se manifestam por meio de dois aspectos fonéticos distintos, o acento lexical (*stress*) e a posição das proeminências (*pitch accents*) (Ladd, 1996a). Os *pitch accents* são eventos locais que assinalam proeminência no nível do enunciado, e estão associados na teoria MA à altura da curva entoacional nos pontos em que esses eventos ocorrem. A literatura sobre o assunto, predominantemente em inglês, destaca a diferença entre *pitch accent* e *stress*, ou seja, entre proeminência e acento lexical. Diferentemente da proeminência, o acento é uma propriedade lexical de sílabas individuais, assinalada no nível da palavra. A teoria MA assume que proeminência e acento lexical são fenômenos distintos, e podem ocorrer alinhados ou não na estrutura métrica do enunciado.

²¹ Para uma melhor adequação com a literatura será usado o termo *pitch accent*, em inglês.

Além da distinção dos *pitch accents* em relação aos acentos lexicais, estes ainda tem relação direta com a altura dos tons. Apesar da teoria MA considerar os *pitch accents* como os fenômenos mais importantes da curva entoacional, estes são analisados à luz da teoria apenas como sequências ou combinações de tons altos e baixos, sendo representados por meio dos rótulos H (*high*) e L (*low*), e pela combinação destes, como aparece na figura 11, que mostra o inventário tonal proposto por Pierrehumbert (1980) adaptado em Botinis et al (2001).

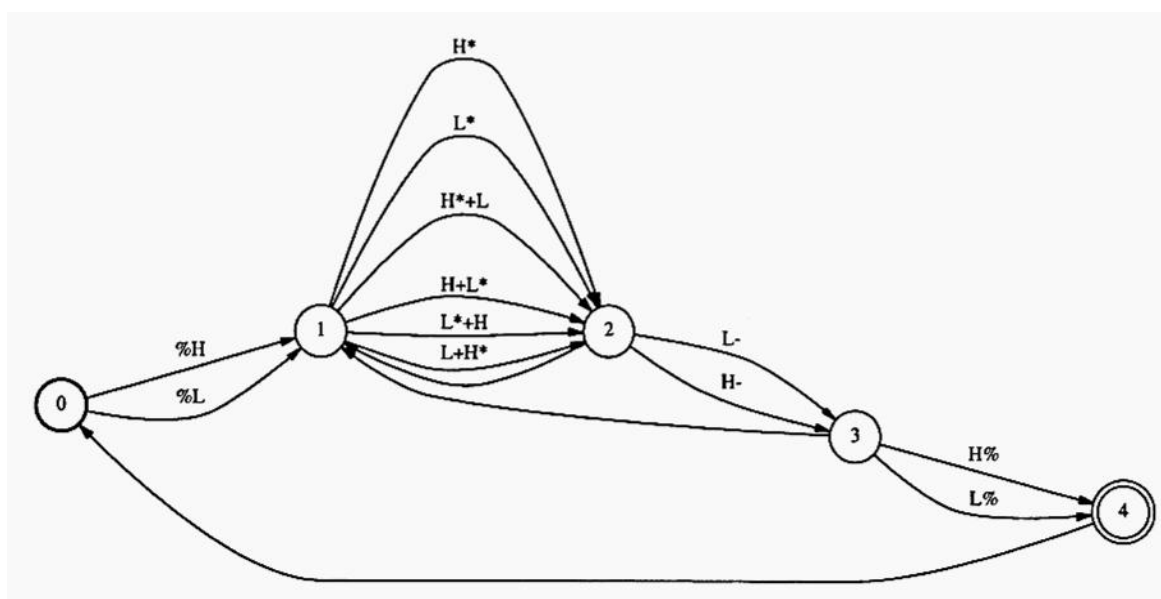


Figura 11: inventário tonal proposto por Pierrehumbert (1980), ilustrando respectivamente *boundary tones*, *pitch accents*, *phrase accents* e *boundary tones*.

Seguindo esses pressupostos, o sistema ToBI (Beckman et al, 1994) apresenta uma sistematização mais econômica da proposta de Pierrehumbert (1980), ainda propondo uma notação que reflete a linearidade da estrutura tonal e a marcação dos *pitch accents*. Tal sistema foi desenvolvido por um grupo de pesquisadores com o intuito inicial de produzir uma ferramenta de transcrição que pudesse auxiliar no entendimento e no desenvolvimento de modelos computacionais de síntese e reconhecimento de fala que requerem grande quantidade de enunciados transcritos prosodicamente (Silverman et al 1992).

Na transcrição do sistema ToBI são utilizadas diferentes camadas, ou *tiers*, para marcações, que servem para a marcação de tons, transcrição ortográfica, agrupamento prosódico e para quaisquer observações sobre a transcrição consideradas importantes para a interpretação de elementos nas outras camadas, como mostra a figura 12.

O elenco de rótulos proposto em ToBI é composto por cinco possibilidades demarcação de *pitch accents*, e de quatro *boudary tones*, que são uma associação entre *phase accents* e altura de fronteiras, como mostra a tabela 3.

Pitch accents:	L*	Tom baixo
	L+H*	Tom baixo seguido de tom alto alinhado a sílaba tônica
	L*+H	Tom baixo alinhado à sílaba tônica seguido de tom alto
	H*	Tom alto
	H+!H*	<i>Downstepping</i>
Boundary Tones:	L-L%	Fronteira baixa precedida por <i>phrase accent</i> baixo
	H-L%	Fronteira baixa precedida por <i>phrase accent</i> alto
	L-H%	Fronteira alta precedida por <i>phrase accent</i> baixo
	H-H%	Fronteira alta precedida por <i>phrase accent</i> alto

Tabela 3: conjunto de rótulos de notação do sistema ToBI

Os *phrase accents* se associam fonologicamente às fronteiras, porém não se alinham foneticamente a nenhum elemento segmental, como a sílaba. Esses elementos possuem uma aceitação controversa nos estudos da entoação, pois marcam proeminência no nível da sentença e sem correlato fonético, o que gera complicações teóricas, como se vê em Grice et al (2000):

Although the phrase accent has gained broad acceptance as a phenomenon relating to the intonation of Swedish, it has been the subject of much debate as far as other languages are concerned. First, it has been argued that the phrase accent constitutes an unnecessary complication to accounts of English and other Germanic languages. Second, even in those analyses that do include a

phrase accent, its exact theoretical status is somewhat unclear. (Grice et al, 2000, p.143).

A notação do agrupamento prosódico das palavras da sentença obedece a uma escala de números que vai de 0 a 4, de forma crescente da sequência mais coesa para a mais disjunta, assim ficam: 0, entre palavras com agrupamento prosódico muito forte, como elisões; 1, entre diferentes palavras prosódicas; 2, para marcar forte disjuntura marcada por pausa ou pausa virtual sem marcas tonais, ou para marcar disjuntura mais fraca do que a indicada por 4, marcando uma fronteira entoacional; 3, para marcar uma fronteira intermediária na sentença; e 4, marcando uma fronteira completa, ou seja, o final de um sintagma entoacional (*intonational phrase*) (Beckmam et al, 1994), como mostra a figura 12 a seguir.

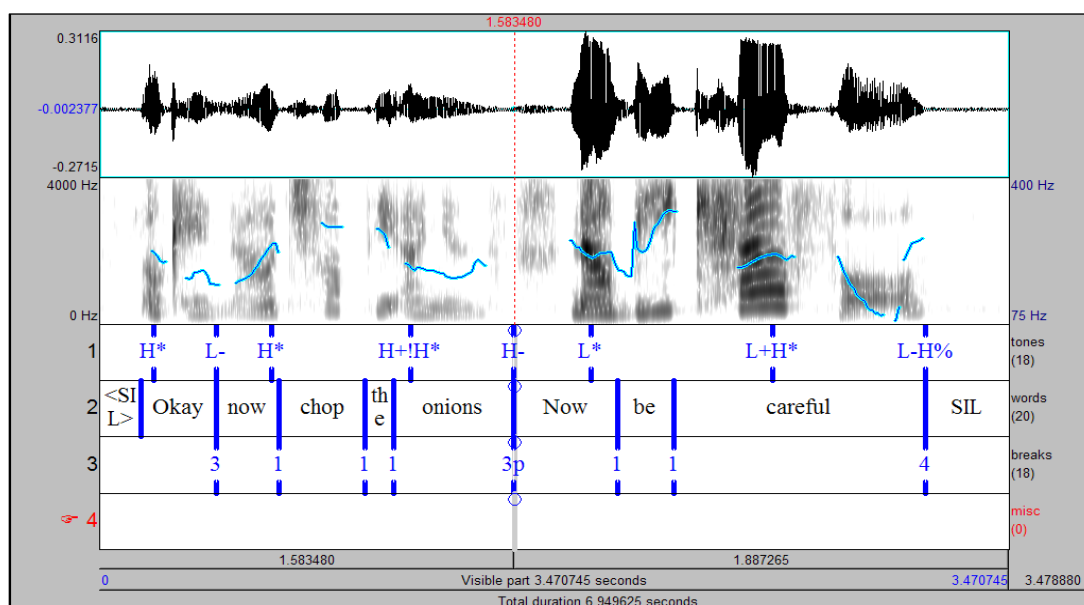


Figura 12: notação ToBI em quatro camadas: *Tones*, *Words*, *Break Indices* e *Miscellaneous*

A implementação desta notação em enunciados específicos é feita por meio de oficinas de notação entoacional. Essas oficinas consistem na reunião de um número de participantes, cujas experiências em relação à notação entoacional sejam de diferentes níveis, para transcrição de arquivos previamente escolhidos. A coincidência de transcrições

entre os participantes é que define o melhor rótulo a ser empregado na representação de determinado tom ou fenômeno entoacional. O sistema ToBI propõe uma coincidência entre transcritores de ao menos 80% (Silverman et al 1992) para o estabelecimento de uma notação. Em Wightman (2002), que coloca em discussão o método de notação ToBI, essa coincidência é de apenas 50%.

Essas oficinas também funcionam para a aprendizagem da forma de transcrição do sistema, que como pressupõe o programa em que se insere, segue um critério que associa informações psicofísicas e físicas a respeito da entoação (Ladd, 1996a). O primeiro passo para a notação nesses sistemas é a percepção de outiva de correlatos como *pitch*, *loudness* e *length* (Lehiste, 1970). A percepção destas características irá determinar a existência de proeminências, fronteiras nos enunciados e movimento de f0, que em seguida se confirma em uma análise dos correlatos físicos presentes na curva entoacional.

1.2 Perspectiva Dinâmica

Enquanto a teoria MA se baseia em aspectos lineares da estrutura tonal, na identificação dos *pitch accents* e no alinhamento abstrato dessa estrutura com o material linguístico, o programa dinâmico-entoacional se concentra na convergência de aspectos fonéticos – velocidade, intensidade, altura, duração – da curva entoacional a fim de atingir um alvo ou desempenhar uma tarefa linguística por meio dos contornos entoacionais, da gama de variação tonal e do alinhamento específico com o material linguístico.

Os contornos entoacionais, enquanto um dos primitivos melódicos desse programa são definidos por uma *trajetória ideal da curva entoacional especificada por um alvo a ser atingido (pitch target) associado a uma unidade segmental, provavelmente uma sílaba* (Xu, 2005). Na especificação dessa trajetória da curva até alcançar seu alvo estão envolvidas, além dos parâmetros físicos, funções comunicativas.

O modelo PENTA (*Parallel Encoding and Target Aproximantion*) proposto por Xu (2005, 2006) ilustra como diferentes funções e parâmetros estão codificados nos

contornos entoacionais. Nessa proposta, funções comunicativas de ordem lexical, frasal, focal, de tópico, etc, são codificadas em parâmetros físicos, como duração, gama de variação tonal e intensidade, que possibilitam o alcance do alvo e seu alinhamento com a sílaba na curva entoacional a uma altura e taxa de aproximação específicas (Xu, 2005; 2006).

Considerando a definição de contorno dinâmico e a proposta do modelo de Xu (2005;2006), a figura 13 ilustra o conceito de contorno dinâmico, que é definido aqui *como uma unidade tonal que contém elementos comunicativos expressos em uma trajetória ideal da curva entoacional, especificada por um alvo a ser atingido e associada a uma unidade segmental linguística*.

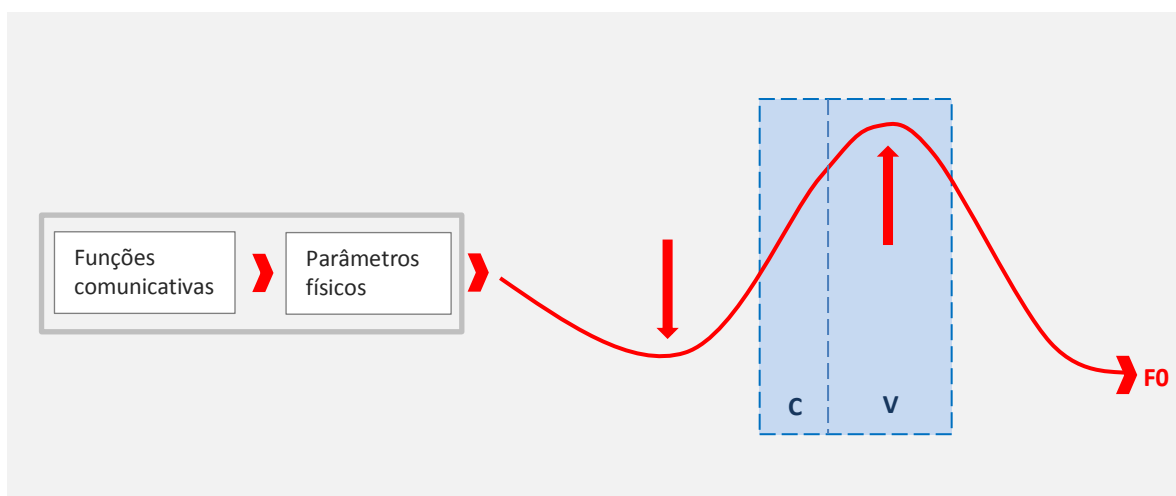


Figura 13: em um contorno entoacional, as (i) funções comunicativas são codificadas em parâmetros físicos; expressas em uma (ii) trajetória ideal de f_0 com um alvo a ser atingido (setas); e (iii) alinhado a uma unidade segmental (sílabas tônicas)

A gama de variação tonal delimita o intervalo de *pitch* no qual um alvo é implementado, e pode ser definido pelos parâmetros do valor do alvo e intervalo de variação. A gama de variação tonal é de grande importância para a especificação de foco, pois é de acordo com as variações do valor do alvo de *pitch* que antecedem e seguem o foco é que é possível percebê-lo. Essas variações de *pitch* podem compor um esquema de codificação do foco, ou mesmo constituírem propriedades intrínsecas a este (Xu, 2005). Da

mesma forma, por meio das variações no valor do alvo de *pitch*, podemos determinar fenômenos como mudanças de tópicos na fala e distinções entre tipos de enunciados, como declarativas e interrogativas. As figuras 14 e 15 ilustram, respectivamente, um enunciado declarativo e um enunciado interrogativo. É possível observar nessas figuras que os padrões de f_0 são distintos e que a gama global de variação no enunciado declarativo é de aproximadamente 90Hz (201-111), enquanto no interrogativo é de aproximadamente 136Hz (233-98).

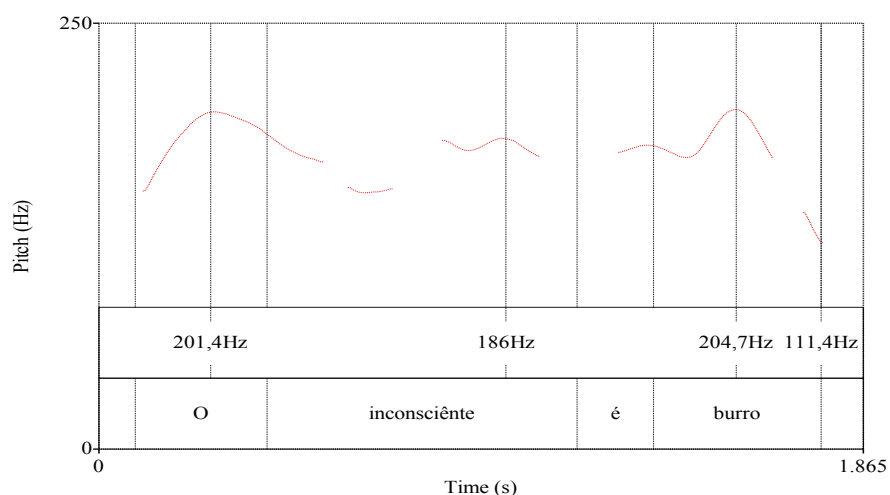


Figura 14: enunciado declarativo com a marcação dos valores dos picos de f_0 em Hertz.

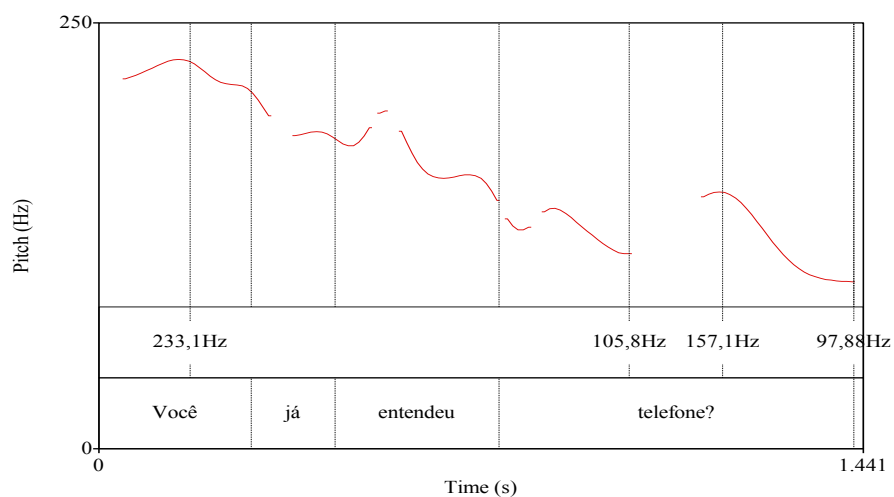


Figura 15: enunciado interrogativo com a marcação dos valores dos picos de f_0 em Hertz.

Cada contorno dinâmico, marcador de proeminência, tem movimento e altura relacionados à altura e ao movimento de f_0 que o antecede e segue, e sua posição na estrutura tonal é determinada pelo alinhamento com o material linguístico, mais precisamente com as sílabas tônicas, ou seja, com o acento. O contorno dinâmico na notação DaTo não se refere à adição de movimentos com diferentes alturas de f_0 , mas a um movimento da curva entoacional ao qual se atribui marcação de proeminência. A figura 16 é ilustrativa de como cada movimento na curva se relaciona com o anterior, pois para cada movimento de subida ou descida na curva, o movimento anterior tem que proporcionar condições para que isso ocorra, ou seja, uma descida para que ocorra subida, e vice-versa.

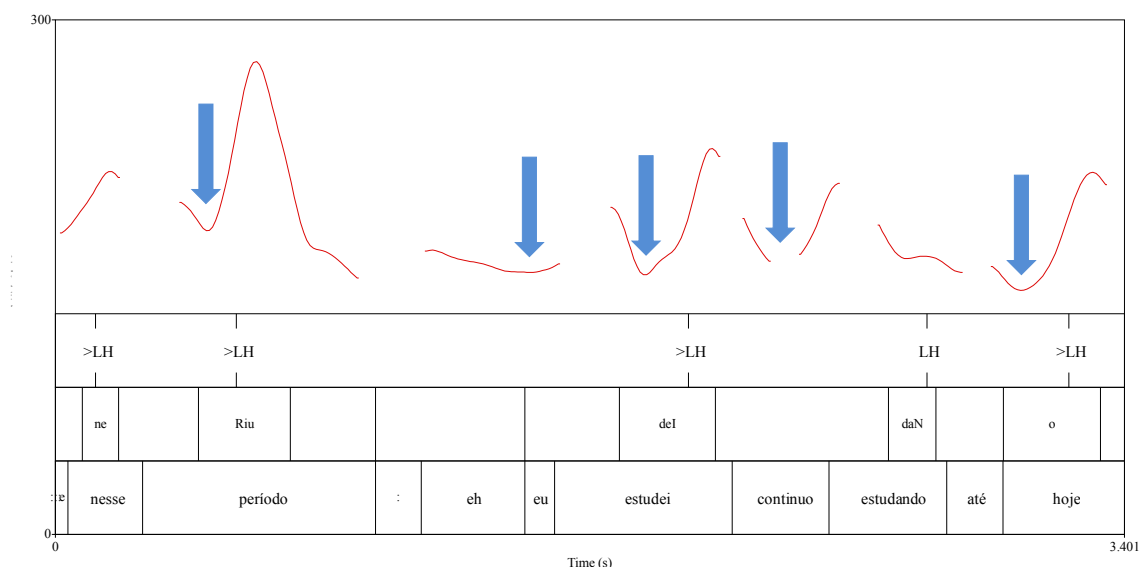


Figura 16: movimentos dinâmicos dos contornos entoacionais com as setas indicando os vales antes das subidas; na primeira camada a notação entoacional, na segunda a marcação das sílabas tônicas, na terceira a transcrição ortográfica.

Essa inter-relação entre os movimentos da curva entoacional defendida pela interpretação dinâmica dos fenômenos que compõem a estrutura entoacional fica evidente ao se extrair a função derivada da curva de f_0 . Ao realizar a extração da função derivada de f_0 por meio do *script F0_extrema.praat* (ver Anexo III) pode-se observar a existência de um controle na produção de f_0 que faz com que esta coordene os movimentos da curva de acordo com a velocidade de subida e descida de f_0 e com o alvo a ser atingido. A figura 17

mostra essa relação: em cinza a curva suavizada de f_0 , em vermelho estão indicados os pontos de mínima altura de f_0 , em verde indicados os pontos de máxima altura. Os pontos que indicam os picos de máxima velocidade de subida (R) e descida (F) de f_0 estão alinhados com a notação na segunda camada da figura, destacados em azul.

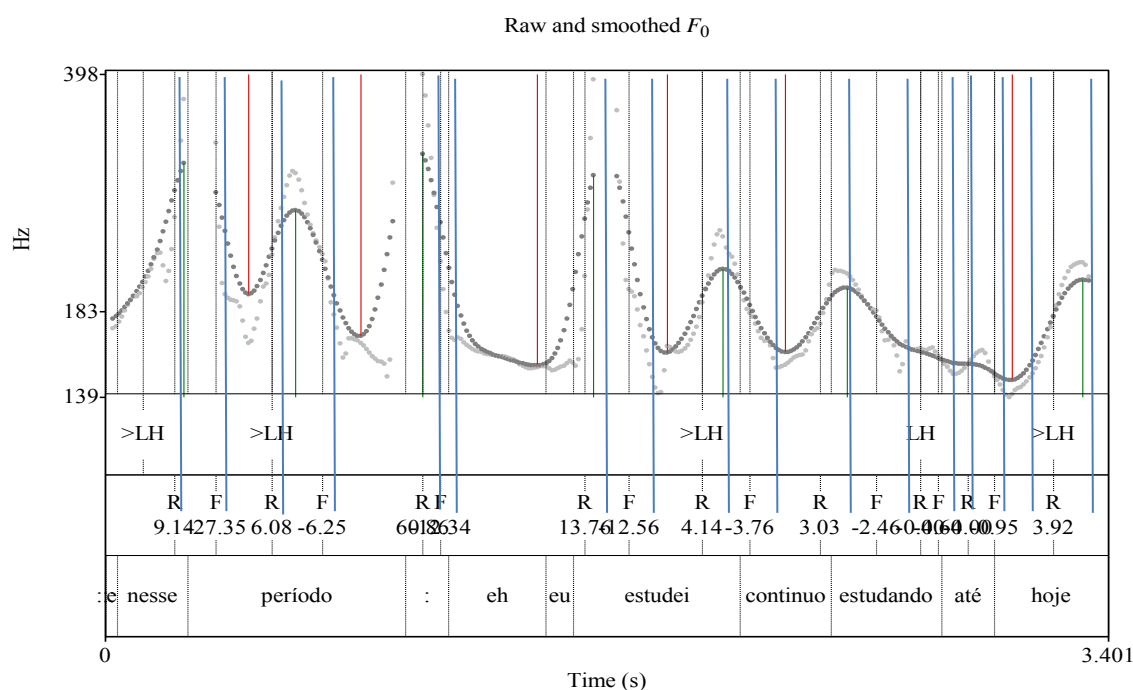


Figura 17: curva entoacional de enunciado; na primeira camada indicando os contornos entoacionais, na segunda os pontos de máxima velocidades de subida e descida de f_0 (linhas azuis) e na terceira a transcrição ortográfica.

O alinhamento específico entre contornos dinâmicos e o material linguístico compreende uma questão central na defesa desta abordagem, pois enquanto a teoria MA estabelece um alinhamento abstrato da curva entoacional cujo ponto específico é decidido por regras de implementação fonética (Pierrehumbert, 1980; 2000), a perspectiva dinâmico-entoacional estabelece um alinhamento específico (Atterer e Ladd, 1994) entre esses elementos, a ponto de este evidenciar a existência de uma sincronia entre os movimentos que produzem f_0 e movimentos articulatorios que geram os padrões espectrais, apesar de esses movimentos serem controlados separadamente. O fato de existir uma sincronia entre movimentos independentes não significa que haja uma total interdependência entre eles,

pois estes podem estar limitados por determinados graus de liberdade na execução de movimentos concomitantes (Xu, 2005).

Os pressupostos e funcionamento da notação DaTo, dentro de uma perspectiva dinâmica da entoação, com contornos dinâmicos e níveis de fronteira, assim como em outros componentes da notação serão apresentados na próxima seção.

1.3 Notação DaTo

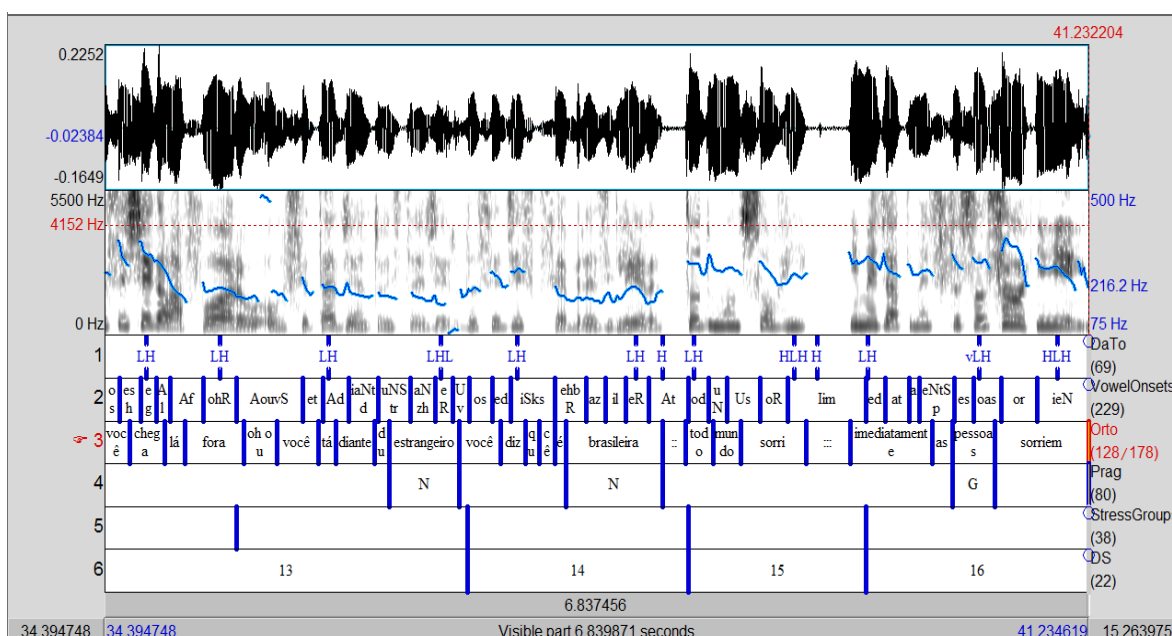


Figura 18: notação DaTo com seis camadas.

O sistema DaTo, além de apresentar uma notação dos fenômenos entoacionais do PB a partir de uma perspectiva dinamicista, procura também apresentar em suas camadas de notação informações pragmático-comunicativas e aspectos fonéticos do ritmo da língua. Para tanto, a notação completa se divide em seis camadas, das quais três são fixas e servem para a notação de: i) contornos dinâmicos, ii) segmentação semi-automática

da fala em unidades V-V, e iii) transcrição ortográfica; e outras três camadas, que podem ser omitidas ou apresentadas de acordo com o objetivo de cada notação, que são: iv) detecção automática de pausas/grupos acentuais, v) informação pragmática e vi) segmentação em unidades discursivas (DS).

A aparência da notação DaTo com todas as camadas no programa Praat (Boersma & Weenink, 2009) pode ser vista na figura 18. A descrição dos contornos entoacionais e das demais camadas de notação do sistema serão apresentadas a seguir.

1.3.1 Contornos Dinâmicos

Os contornos dinâmicos, como definido anteriormente, são movimentos contínuos de f_0 , que percorrem a transição de um ponto a outro na curva entoacional até atingir o seu alvo. Este movimento de transição carrega a informação tonal que antecede e segue seu alvo. Nessa concepção, um contorno dinâmico torna relevantes as transições até então consideradas pela teoria MA como apenas transições preenchedoras entre eventos estáticos.

No sistema DaTo, os contornos dinâmicos se dividem entre ascendentes, descendentes e níveis de fronteira, e são representados por rótulos que ilustram o movimento de f_0 pelas letras H (*high*) e L (*low*), assim com na teoria MA, esses rótulos foram mantidos por julgar tais rótulos produtivos e ilustrativos, e, portanto, de mais fácil compreensão.

O conjunto de contornos entoacionais do sistema DaTo são anotados levando em consideração seu padrão de movimento e alinhamento com as sílabas tônicas partindo primeiramente da percepção de ênfase e posteriormente de uma inspeção visual da curva entoacional que determinará o rótulo a ser empregado. Diferentemente, a teoria MA pressupõe inicialmente a percepção de outiva dos *pitch accents* e posteriormente sua verificação na curva. Um dos pressupostos do treinamento para notação nos sistemas que seguem a teoria MA é a concordância entre sujeitos sobre o tipo de movimento tonal que conseguem detectar de outiva, sem a inspeção na curva entoacional.

1.3.1.1 Teste de Percepção: contornos ascendentes e descendentes

Para avaliar a dificuldade em atingir a concordância entre sujeitos sobre o tipo de movimento tonal – desnecessária no sistema DaTo – que indica proeminências, foi proposto a um grupo de 11 sujeitos um teste MFC no programa Praat, por meio de um *script* desenvolvido especificamente para tal teste (ver Anexo IV). O teste consistia em que os sujeitos respondessem com base em 10 estímulos sonoros transcritos simultaneamente na tela do computador, se o contorno das palavras selecionadas (em caixa alta) nestes enunciados, era ascendente, descendente ou se eles não conseguiam definir. Ao todo eram apresentados 17 estímulos/sentenças, sendo sete destas repetições com uma palavra diferente destacada. Cada um dos 17 estímulos/sentenças era repetido duas vezes, somando assim 34 estímulos. Foram selecionados para o experimento oito linguistas familiarizados com notação entoacional e três músicos, dos quais cinco homens seis mulheres, na faixa etária entre 25 e 40 anos.

O teste Kappa Fleiss de concordância entre as escolhas dos sujeitos feito no programa R, obteve resultado $k = 0,19$ e $z = 10,1$ ($p < 10^{-7}$), significativo para concordância baixa. Além da concordância baixa na percepção dos tipos de contornos, os sujeitos não foram precisos em avaliar as repetições, pois os sujeitos, sem exceção, atribuíram em algum momentos em estímulos, um contorno à primeira repetição e outro diferente à segunda.

O resultado do teste tem resultado contrário ao que propõe Escudero et al (2012), que encontra uma taxa significativa de concordância entre transcritores quando as taxas de discordância são transformadas em distâncias métricas possibilitando o uso de uma escala multidimensional para visualizar a discordância entre transcritores. Porém, vai de encontro ao que Hist (2005) defende quando sugere que os transcritores, ou ouvintes, estão mais atentos à função prosódica do que à forma. Ou seja, os sujeitos, mesmo com conhecimento entoacional, não são precisos na percepção dos tipos de contornos entoacionais, a não ser que disponham da observação da curva entoacional, ou seja, do material fonético. Esse é mais um ponto a favor de um sistema dinâmico, que pressupõe a interação entre componentes fonológicos e fonéticos para notação entoacional.

1.3.1.2 Contornos Ascendentes: LH, >LH, vLH e HLH

Os contornos ascendentes são os mais frequentes na entoação do PB e estão constantemente associados à marcação de ênfase e foco estreito. Estes contornos têm como referência para o alinhamento da curva com a sílaba tônica da palavra focada a posição alta (H) do contorno, e apresentam um padrão de movimento determinado por características acentuais que antecedem e sucedem o foco.

Antecedendo o foco, estes contornos apresentam uma descida de f0 durante as sílabas pretônicas, considerada obrigatória para a realização do movimento posterior, que é a subida de f0 para a realização do movimento ascendente. Em posição pós-tônica verifica-se para todos os contornos ascendentes um *resetting*²². Esse *resetting* mostra que a percepção do foco está associada, além do emprego de maior energia na sua realização, à oposição ocorrida com as sílabas não acentuadas posteriores ao foco. Estas características mostram estratégias linguísticas que envolvem todo o sistema de produção e percepção de ênfase.

i) LH - *rising contour*

O contorno dinâmico LH ou *rising* é bastante produtivo na marcação de ênfase e foco estreito²³ em PB e ocorre em posições iniciais e mediais de sentenças declarativas e interrogativas. Seu movimento consiste em (i) uma subida do contorno que se inicia na sílaba pretônica imediatamente anterior à tônica, (ii) culmina em um pico alinhado à vogal da sílaba tônica da palavra focada, e (iii) termina com a descida da curva na sílaba pós-tônica. Um exemplo do contorno dinâmico LH está na figura 19, alinhado à vogal tônica da palavra “globo”.

²² Com exceção do contorno vLH, como se verá a seguir.

²³ Embora caminhando juntos, a ênfase se refere ao fenômeno fonético enquanto o foco se refere ao fenômeno sintático. Vamos usar aqui o termo *ênfase*.

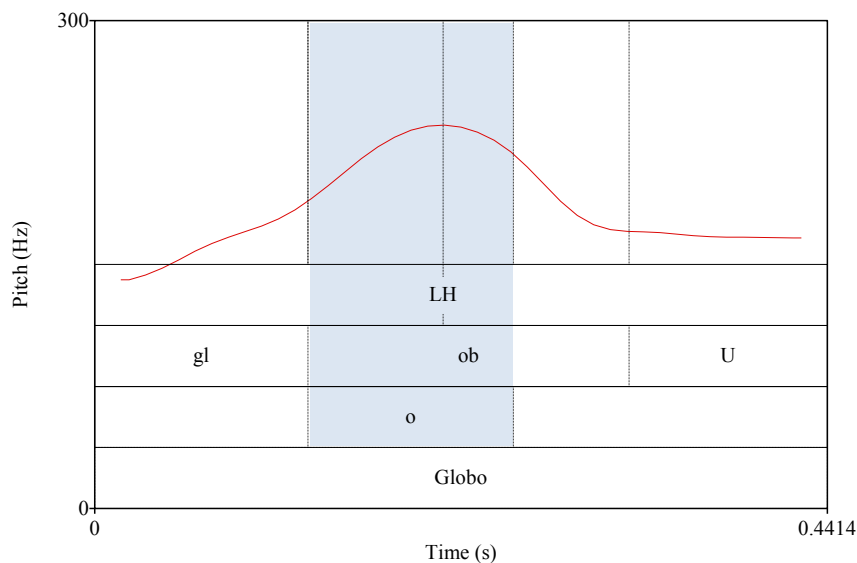


Figura 19: contorno LH alinhado à vogal tônica da palavra “globo”.

ii) *>LH - late rising*

O contorno dinâmico *>LH*, ou *late rising*, possui as mesmas características formais do movimento que compõe LH, ocorrendo nas mesmas posições nos enunciados. No entanto, o alinhamento do pico com a vogal tônica da palavra alvo ocorre com atraso. Em *>LH* a vogal tônica se alinha em toda sua extensão à subida de f_0 (figura 20). Esse desalinhamento entre a curva entoacional e o material linguístico evidencia uma estratégia funcional para delimitação de ênfase que se utiliza do resultado de uma perturbação no alinhamento entre os sistemas articulatório e entoacional.

>LH, assim com LH, ocorre com bastante frequência na marcação do movimento inicial de enunciados declarativos e interrogativos, como se podem ver assinalados na figura 21, em que *>LH* ocorre no início do enunciado e LH após a uma fronteira entoacional (marcada por H). Esse recurso, como se verá mais adiante neste texto, tem função de indicar o início de novos segmentos discursivos.

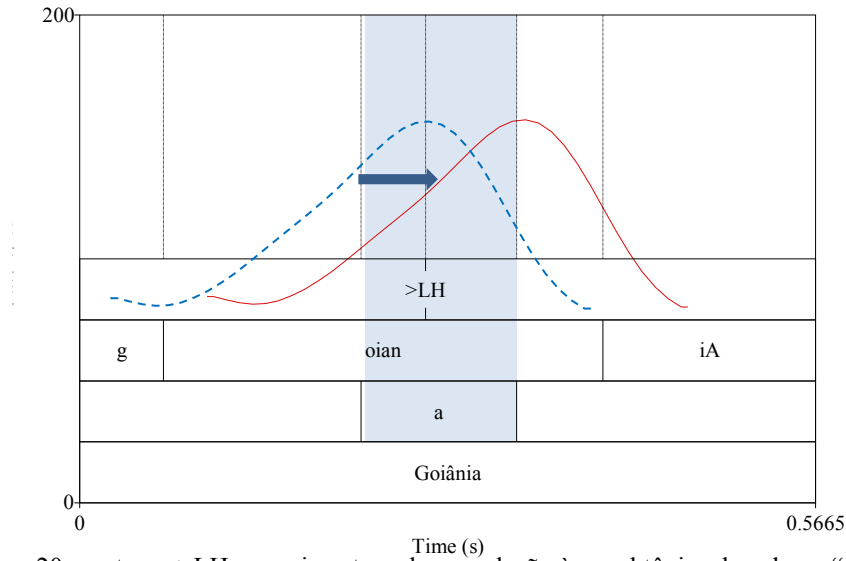


Figura 20: contorno >LH com pico atrasado em relação à vogal tônica da palavra “Goiânia”

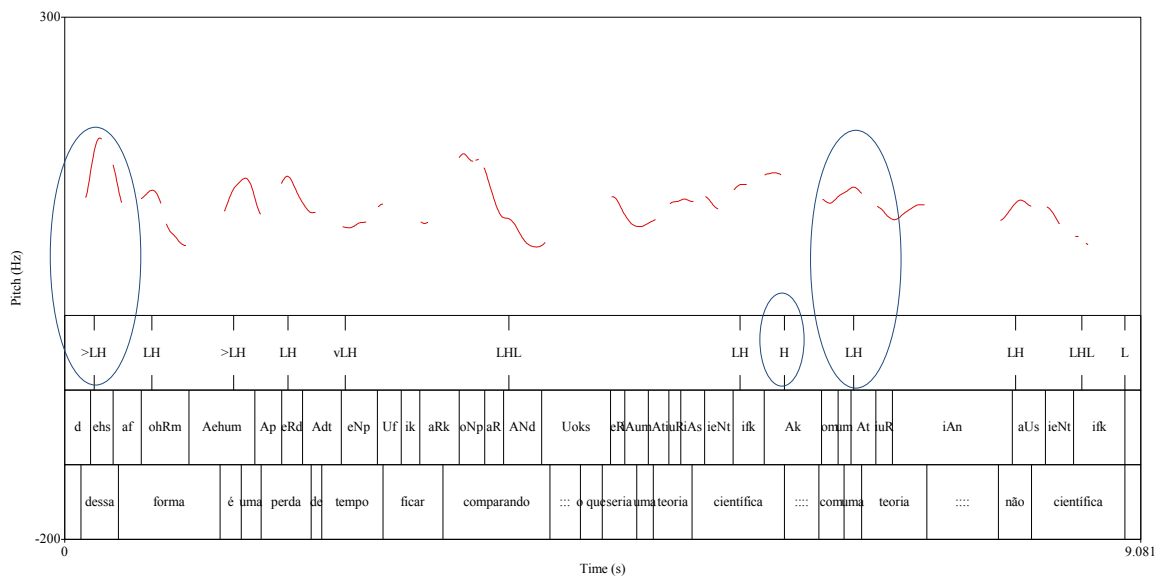


Figura 21: >LH na marcação do movimento inicial f0 e LH marcado após fronteira H

iii) Experimento de Percepção LH, >LH: alinhamento e altura

Os contornos entoacionais LH e >LH possuem o mesmo padrão dinâmico de movimento e aparecem em mesma proporção de ocorrências na notação entoacional do PB,

com predominância de >LH em posição inicial de enunciados declarativos. No entanto, a diferença de alinhamento entre estes contornos, que justifica a atribuição de diferentes rótulos a eles, pode indicar diferenças funcionais entre ambos. A observação e a notação destes apontam para o contorno >LH como marcador de proeminências mais fortes em relação às marcadas por LH no PB. Para verificar a validade desta observação foi montado um experimento de percepção que se baseou nas observações preliminares sobre a relação entre alinhamento, altura da curva e grau de ênfase (Kohler, 2005).

Para a realização dos testes foram feitas manipulações mecânicas na curva entoacional de um conjunto de enunciados. Os enunciados que apresentavam originalmente contornos LH foram manipulados mecanicamente para apresentarem contornos >LH, e vice-versa. A manipulação desses contornos foi feita no programa Praat por meio da ferramenta *manipulation*. Esta ferramenta possibilita fazer a estilização da curva de f0 nos enunciados a partir de quantidade de pontos especificada, e apresenta a possibilidade de ressíntese do sinal de fala contendo a curva de f0 manipulada.

A primeira manipulação realizada foi no alinhamento dos contornos originais, sem alteração na altura. Para cada arquivo de som foram feitas manipulações no alinhamento, nas quais os deslocamentos da curva em relação à posição central da vogal tônica foram de 50ms e 100 ms. Dessa forma, para contornos LH, em que o pico de f0 se alinha à vogal tônica, este ficou com atraso de 50 e 100ms em relação ao pico original, e para contornos >LH, em que o pico de f0 ocorre após a vogal tônica, o alinhamento do pico ficou adiantado em 50 e 100ms. A figura 22 mostra a curva original em vermelho, com contorno >LH, e as manipulações em 50ms (manipulação1 – M1, em verde) e 100ms (manipulação 2 – M2, em azul).

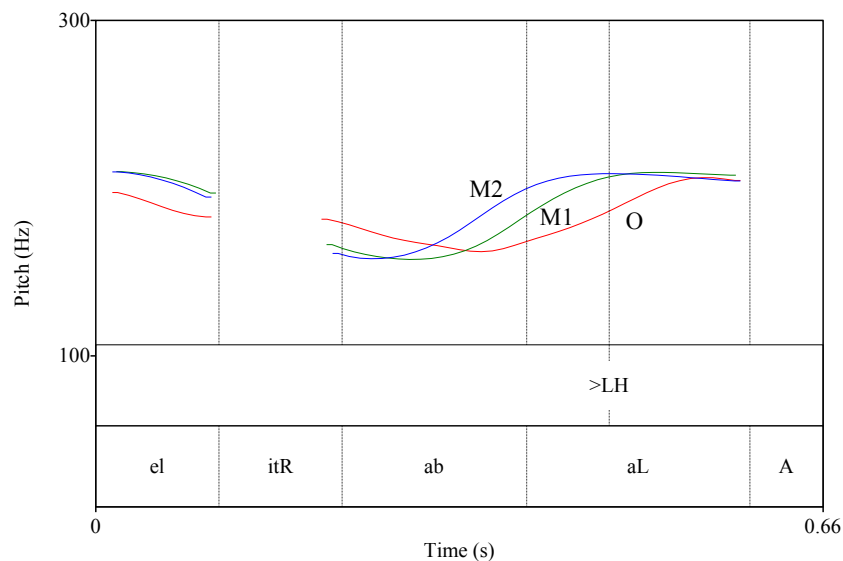


Figura 22: curva original da palavra “trabalha” em vermelho com contorno >LH, e as manipulações em 50ms (verde) e 100ms (azul).

A segunda manipulação foi quanto à altura dos contornos. Com o intuito de avaliar se diferentes alturas produzem algum efeito sobre a percepção da ênfase, pois assim como o alinhamento produz em >LH sensação de mais ênfase, a altura também parece ser um fator que determina essa diferença. Portanto, para os contornos originais foram produzidos dois novos contornos manipulados, com acréscimo de aproximadamente 1,5 semitons (manipulação 1 – M1) e 3,0 semitons (manipulação 2 – M2) respectivamente, como se pode ver na figura 23. A escolha por medidas em semitons e não em *Hertz* para o acréscimo na altura dos picos se deve ao fato de que o semitom, por ser uma medida psicoacústica, possibilita melhores resultados em testes de percepção quando se faz a relação entre produção e percepção (Rietveld & Gussenhoven, 1988)

Uma terceira manipulação foi realizada somando as manipulações na altura e no alinhamento. Portanto, para cada contorno original que apresentava manipulação no alinhamento em 50ms (M1) e 100ms (M2), foram feitas manipulações nas alturas destes, acrescentando 2st na altura de cada uma das curvas já manipuladas em relação ao alinhamento, gerando mais duas curvas (M3 e M4), como mostra a figura 24.

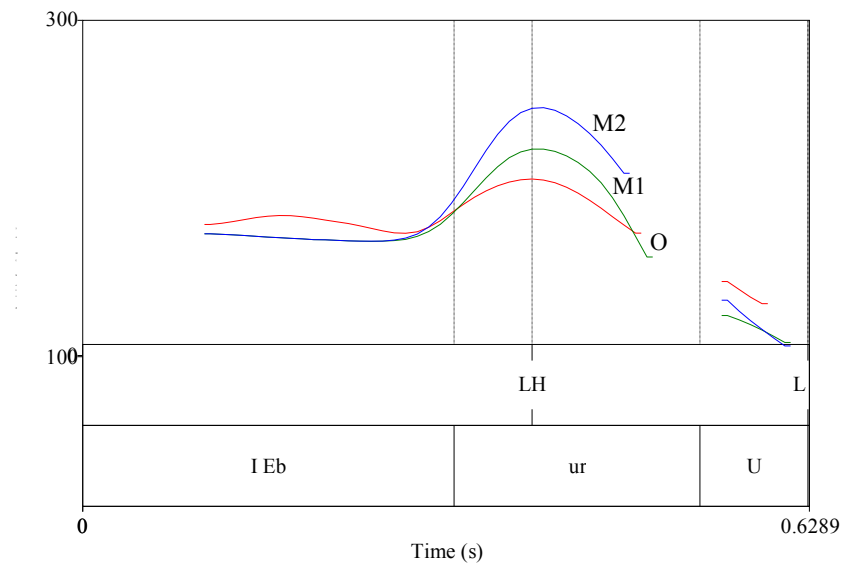


Figura 23: curva original da palavra “burro” em vermelho e as manipulações na altura em 1,6 st (verde) e 3,0 st (azul).

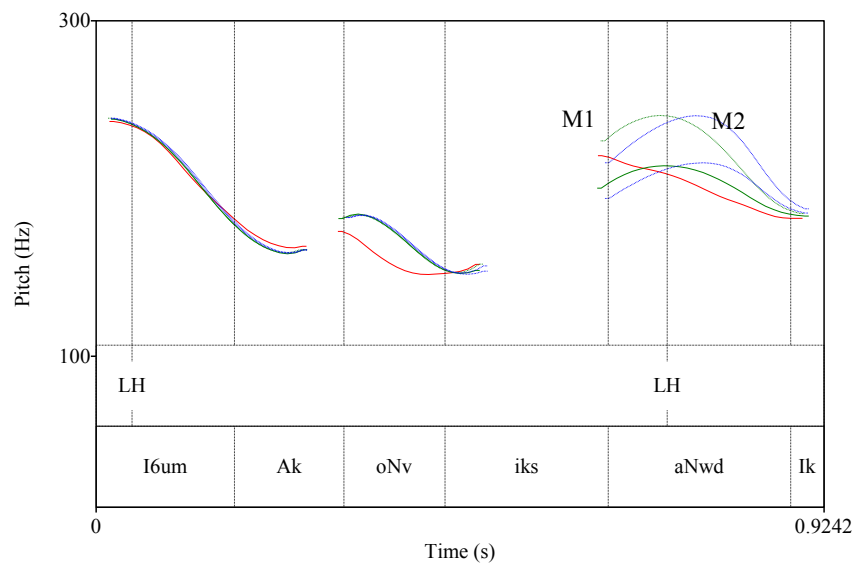


Figura 24: curva original da palavra “convicção” em vermelho, com contorno >LH, e as manipulações em 50ms (verde), 100ms (azul), acréscimo de 2st em relação a M1 (verde pontilhado) e acréscimo de 2st em relação a M2 (azul pontilhado).

Um conjunto de três testes do tipo MFC (usando *scripts* MFC no Praat como o apresentado no anexo IV) foi realizado com a finalidade de comparar cada um dos conjuntos de manipulações (M) com as gravações originais (O) e também entre as próprias manipulações (cf. Lucente & Barbosa, 2010). Sendo assim, os dois primeiros testes avaliaram a percepção de ênfase entre os contornos manipulados, um quanto ao alinhamento e outro quanto à altura de acordo com seguinte combinação de estímulos:

$$\begin{array}{l} O - M^1 \\ O - M^2 \\ M^1 - M^2 \end{array}$$

O terceiro teste, que avaliava o grau de ênfase nas curvas manipuladas quanto à altura e altura/alinhamento em relação às gravações originais seguiu a seguinte combinação de estímulos:

$$\begin{array}{l} O - M^1 \\ O - M^2 \\ O - M^3 \\ O - M^4 \end{array}$$

Esse conjunto de testes apresentou os seguintes resultados: o primeiro teste, para mudanças no alinhamento de f_0 , teve como resultado para concordância entre sujeitos coeficiente $k=0,12$, e $z=9,44$, resultado significativo para concordância baixa para >LH mais proeminente, o que mostra que os sujeitos não percebem >LH necessariamente como mais proeminente que LH. Com base no alinhamento, o segundo teste, para mudanças na altura de f_0 , teve como resultado para concordância entre sujeitos coeficiente $k=0,04$, e $z=2,66$, resultado pouco significativo e concordância muito baixa, o que mostra que os sujeitos não são coesos na percepção de diferenças de proeminência a partir de mudanças na altura de f_0 , sejam estas proeminências mais ou menos acentuadas. Por fim, o terceiro teste, que avaliou em conjunto as mudanças de alinhamento e altura de f_0 , teve como resultado para

concordância entre sujeitos coeficiente $k=0,16$ e $z=11,9$, resultado bastante significativo, porém, com um índice de concordância entre sujeitos ainda baixo.

Esses resultados levaram à conclusão de que mudanças no alinhamento e na altura de >LH em relação a LH são baixas quanto a diferenças funcionais. Os resultados sugerem que a diferença de alinhamento entre esses contornos se deve a uma restrição temporal para alcançar o alvo do contorno ascendente, o que faz com que o sistema se reajuste para tal tarefa promovendo um desalinhamento entre o f0 e o sistema articulatório. Apesar de não clara a diferença funcional entre LH e >LH, o contorno >LH fica reservado na maioria das ocorrências à subida inicial dos enunciados declarativos, sugerindo uma função marcadora de mudança de DS (ver mais detalhes no capítulo Interação Ritmo-Discurso-Entoação).

iv) vLH - compressed rising

O contorno vLH, chamado por *compressed rising*, possui a mesma capacidade de LH e >LH de marcar ênfase. Esse contorno ocorre alinhado à uma vogal tônica entre dois picos de f0, em uma posição em que não há espaço, ou possibilidade fisiológica, para a realização de outro pico de valor elevado, por isso é chamado de LH comprimido, como se pode ver na figura 25. Em casos como estes, em que existe a necessidade de uma ênfase em determinada parte do enunciado, porém aspectos fonéticos impedem que haja a realização de um contorno LH, o sistema de produção da fala consegue solucionar o problema realizando a ênfase em um contorno comprimido, mas que perceptualmente tem o mesmo efeito dos contornos LH e >LH.

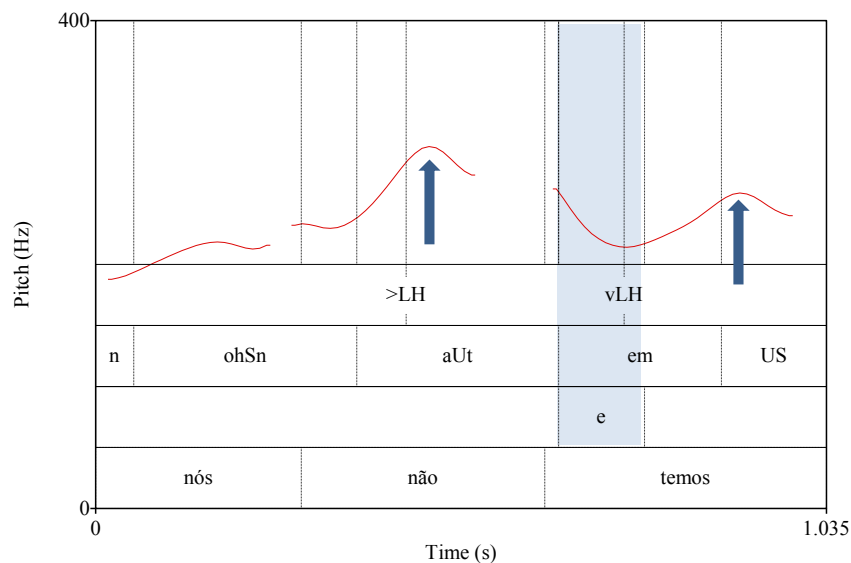


Figura 25: vLH alinhado a vogal tônica da palavra “temos” entre dois picos de f_0 .

Essa propriedade do sistema entoacional do PB de realizar um contorno comprimido está diretamente relacionada às restrições do movimento de f_0 , que precisa realizar um movimento de subida ou descida que possibilite que a curva alcance seu próximo alvo. A figura 26 mostra a relação da curva entoacional com os pontos de máxima velocidade de descida e subida de f_0 . É possível observar nessa figura que o contorno vLH se encontra entre o ponto de máxima velocidade de descida de f_0 (F_{max}) e o ponto de máxima velocidade de subida de f_0 (R_{max}) (posição na qual seria impossível a realização de um pico alinhado à vogal tônica), e irá formar o pico alinhado à vogal postônica da palavra “temos”. O movimento de f_0 na figura 26 mostra ainda, de forma bastante dinâmica, que o pico que se realiza após o contorno vLH é de menor altura em relação ao que antecede vLH, pois para que o segundo pico atingisse altura maior seria necessário mais tempo (indicado pela linha azul pontilhada), e isto faria com que o pico se alinhasse à palavra “lugar”, mudando assim a intenção de ênfase do falante.

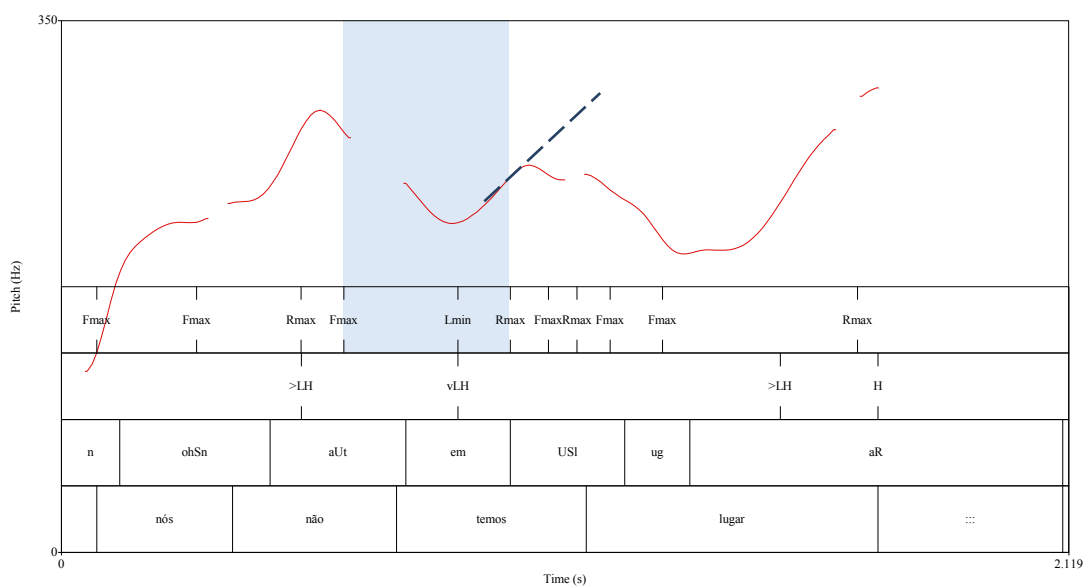


Figura 26: vLH entre os pontos de máxima velocidade de descida e de subida de f_0 .

v) *HLH* - *falling-rising* ou “*contorno do camelo*”

O contorno entoacional descrito pela notação HLH compreende a realização de dois picos sucessivos de f_0 , sendo que o primeiro se realiza alinhado a uma sílaba pretônica, geralmente a primeira sílaba da palavra, realizando uma proeminência. O segundo pico ocorre alinhado à sílaba tônica, seguindo o mesmo padrão de movimento e alinhamento dos contornos LH ou >LH, o que confirma a robustez do padrão dos contornos ascendentes. Não existe um padrão de altura hierárquico para estes picos podendo a proeminência que ocorre na sílaba pretônica ser realizada com maior gama de variação tonal do que a que ocorre na sílaba tônica, e vice-versa. Esta sequência de dois picos fez com que esse contorno fosse apelidado de “contorno do camelo”.

A observação deste contorno mostra que se trata da realização do mesmo tipo de foco, porém, com uma ênfase maior auxiliada pela realização de uma proeminência na primeira sílaba, que produz um efeito amplificador aos contornos LH e >LH, e por isso pode ser acompanhado pelo símbolo “>”, indicando atraso. A realização deste tipo de

contorno está relacionada à realização de focos, no qual a palavra focada deve receber mais destaque no contexto em que é produzida a sentença.

Este tipo de contorno aparece em palavras com diferentes números de sílabas em posições e sequências diversas, como se vê em Lucente (2008, p. 39). Na figura 27, por exemplo, o contorno HLH aparece em uma palavra de cinco sílabas, e mantém o padrão descrito para esse contorno, que é um pico na primeira sílaba, uma descida obrigatória e o segundo pico na sílaba tônica, alinhado ou em atraso em relação à vogal tônica.

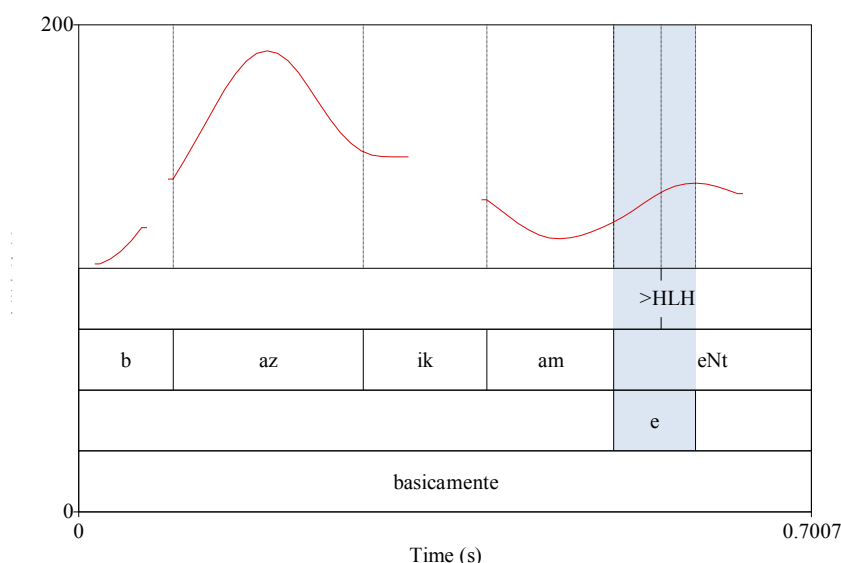


Figura 27: contorno HLH em palavra “basicamente”: pico na primeira sílaba (ba), uma descida obrigatória nas sílabas pretônicas e o segundo pico na sílaba tônica (men), em atraso em relação à vogal tônica.

1.3.1.2 Contornos descendentes LHL, HL, >HL e vHL

Os contornos descendentes seguem a mesma perspectiva dos ascendentes, porém tendo como referência para o alinhamento com a sílaba tônica a posição baixa (L). O movimento *default* de f_0 para estes contornos se inicia após uma subida suave de f_0 , que pode anteceder determinado contorno ou mesmo estar presente dentro dos limites deste, ocorrendo em seguida uma queda que se alinha à sílaba tônica da palavra focada. O *resetting* das sílabas pós-tônicas ocorre da mesma forma que nos contornos ascendentes, confirmando uma estratégia do PB para realização de ênfase.

i) *LHL- rising-falling*

O contorno descendente LHL representa na notação DaTo o movimento classificado como *dowstepping* pela teoria MA. Este contorno descreve um *resetting* de f0 próprio de finais de enunciados declarativos, como descrito por Moraes (1998). Em LHL ocorre, em oposição aos contornos ascendentes, uma subida obrigatória de f0, que pode ser um contorno ascendente que o antecede, realizando assim um movimento de subida e descida mais amplo do que o que ocorre em LH, por exemplo. Esta subida que inicia o movimento de *dowstepping* pode ocorrer na última sílaba da palavra anterior – ou mesmo na partícula funcional anterior à palavra focada –, na sílaba tônica da palavra anterior, assim como na primeira sílaba da palavra alinhada ao contorno, se essa apresentar número de sílabas que possibilite espaço para uma descida suave formando sucessivos degraus até alcançar a sílaba tônica.

A figura 28 ilustra esse movimento, que se inicia na descida de f0 após a realização do contorno >LH antecedente e culmina na vogal tônica da palavra “acabar”.

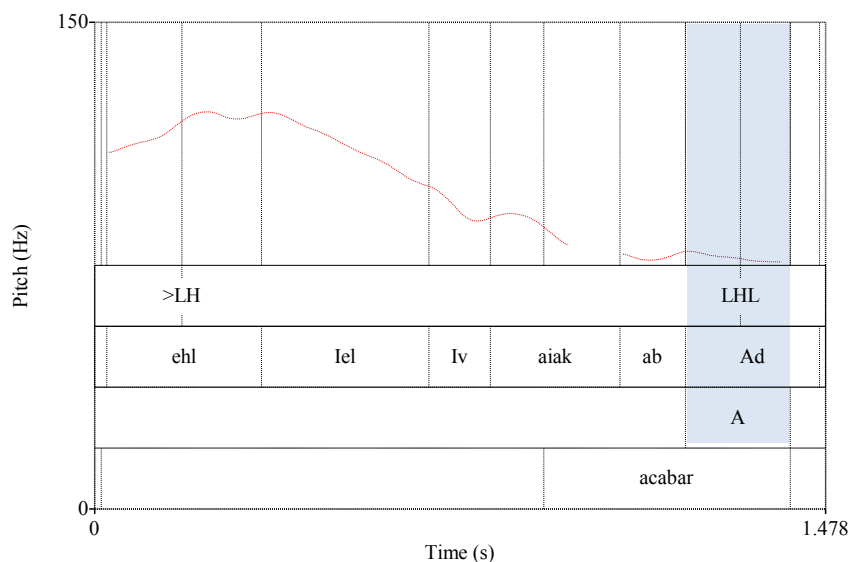


Figura 28: contorno LHL, que se inicia no movimento de descida do contorno >LH e culmina na vogal tônica da palavra “acabar”.

ii) *HL - falling*

O contorno dinâmico HL, encontrado frequentemente em fronteiras finais de sentenças declarativas, talvez seja o contorno mais complexo em termos de movimento, que é composto por 4 estágios: i) a subida obrigatória de f_0 na sílaba pretônica, alinhada ao início da consoante; ii) pico de f_0 alinhado ao meio da vogal pretônica; iii) descida iniciada após o pico coincidindo com a transição da vogal pretônica para a consoante da sílaba tônica; iv) vogal tônica ocorrendo na mesma altura que a fronteira L. As sílabas pós-tônicas, quando existem, também estão no nível da fronteira.

Para palavras que apresentam a primeira sílaba tônica, o padrão encontrado é o mesmo, porém com a subida obrigatória ocorrendo na última sílaba da palavra anterior, mostrando a necessidade desta subida para a realização de ênfase. A figura 29 mostra o contorno HL e cada etapa de seu movimento, desde a subida inicial até seu alinhamento com a vogal tônica da palavra “manhã”.

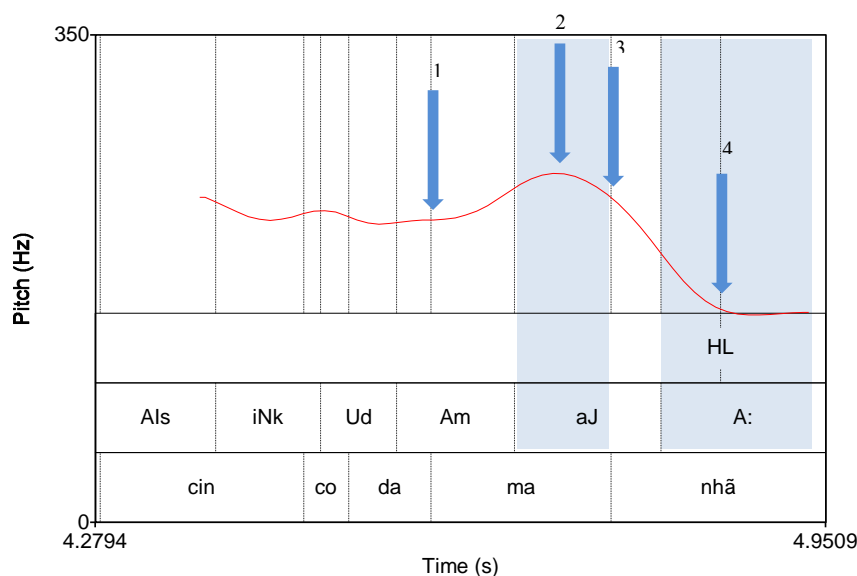


Figura 29: contorno HL e as quatro etapas seu movimento

iii) Experimento de percepção HL, LHL: alinhamento e altura

Os contornos dinâmicos descendentes LHL e HL aparecem ambos em fronteiras finais de enunciados declarativos no PB. Porém, por se tratarem de contornos distintos com movimento e alinhamento distintos, um experimento de percepção foi aplicado a fim de verificar a existência de diferenças funcionais entre os dois contornos.

Da mesma forma como foi feito para o experimento com os contornos LH e >LH, para a comparação e julgamento das diferenças entre LHL e HL um conjunto de enunciados declarativos que apresentam contorno final LHL foram manipulados manualmente no programa Praat para apresentarem contornos HL na posição em que anteriormente se apresentavam os contornos LHL (Lucente e Barbosa, 2010).

A manipulação do alinhamento consistiu no deslocamento do pico de f_0 , que em LHL se alinha à vogal tônica das palavras anteriores à palavra marcada por LHL, para a vogal tônica da última palavra do enunciado, como ocorre em HL, fazendo com que a descida de f_0 fique comprimida em um espaço menor de tempo. Quanto à manipulação na altura dos picos, estes foram acrescidos em três semitons, no intuito de verificar se a altura dos picos influencia na percepção das funções dos contornos, assim como foi feito com LH e >LH.

Um exemplo das duas manipulações feitas no contorno LHL aparecem na figura 30: em vermelho aparece a curva original, como notação LHL, pico alinhado à sílaba tônica de “mundo” e fronteira final L; em azul aparece a curva manipulada, com o pico alinhado à sílaba tônica de “acabar”, alterando a notação de LHL para HL, com um *pitch range* mais comprimido em relação ao contorno original e fronteira L; em verde aparece a manipulação na altura do pico, em que este é acrescido em três semitons, fazendo com que a descida de f_0 seja ainda mais acentuada.

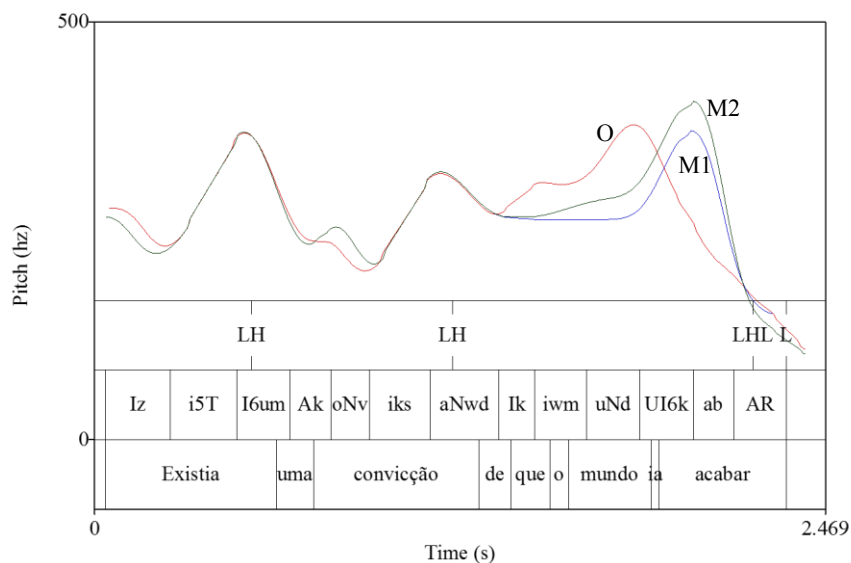


Figura 30: manipulações feitas no contorno LHL: curva original (em vermelho), como notação LHL; curva manipulada alterando contorno de LHL para HL (em azul); manipulação na altura de f0 (em verde)

Antes da realização do experimento MFC foi feito um pré-experimento de escolha livre para verificar que tipo de função é atribuída a cada um dos contornos. Esse pré-experimento consistia na apresentação de 24 pares de enunciados aos sujeitos, compostos pela combinação de 12 enunciados originais (O) com 24 manipulações, sendo 12 em relação ao alinhamento (M1) e 12 em relação à altura (M2). Esses pares de enunciados foram apresentados em fones-de-ouvido para 10 sujeitos, sete mulheres e três homens, todos estudantes universitários na faixa etária dos 20 anos. A combinação dos pares apresentados consistia na reprodução do enunciado original seguindo de uma das manipulações, por exemplo:

O – M1

O – M2

A tarefa dos sujeitos era apontar livremente se eles identificavam algum tipo de diferença no segundo enunciado do par em relação ao primeiro, e que tipo de diferença era essa. As respostas à tarefa foram bastante distintas, mas dentre as respostas dos sujeitos que identificaram a existência de uma diferença entre os pares, as mais comuns foram: mais

enfática, mais explicativa e mais finalizadora. Pela noção de finalização apontada pelos sujeitos foi inferido que se tratavam de enunciados afirmativos ou conclusivos. Usando os parâmetros de ênfase e finalização apontados pelos sujeitos, um teste foi montado para avaliar se as diferenças comunicativas e funcionais entre os contornos LHL e HL eram baseadas em diferenças de ênfase e finalização dos enunciados.

O experimento em si consistiu em três testes MFC, um para avaliar diferenças de ênfase e outro para avaliar diferenças na finalização dos enunciados. Para estes dois testes foram usados os mesmos *inputs*, no entanto esses formaram 36 pares de enunciados, compostos pelos originais em pares com as manipulações e as estas formando pares entre si, da seguinte forma:

O – M1

O – M2

M1 – M2

Esse experimento foi aplicado a 15 sujeitos distintos dos sujeitos que realizaram o pré-experimento, sendo estes nove mulheres, seis homens, todos estudantes universitários na faixa etária entre 20 e 30 anos.

Ambos os testes MFC foram aplicados no programa Praat, por meio de três diferentes *scripts*²⁴. Nesse experimento os sujeitos ouviam os pares de enunciados através de um fone-de-ouvido ao mesmo tempo em que na tela do computador os enunciados eram apresentados na forma escrita. Na primeira tarefa, ao final da apresentação dos pares, os sujeitos eram questionados se a segunda repetição era mais enfática, menos enfática ou igual à primeira repetição e o sujeito deveria responder clicando na tela sobre a resposta que julgasse correta. Na segunda tarefa, após a apresentação de cada par, os sujeitos deveriam julgar se a segunda repetição era mais finalizadora ou menos finalizadora em relação ao primeiro enunciado.

O terceiro teste de escolha forçada foi realizado para verificar se o léxico influenciava na decisão sobre a finalização dos enunciados. Para isso um novo *script* foi aplicado a 14 novos sujeitos, sendo 10 mulheres e 4 homens, estudantes universitários,

²⁴ Assim como o *script* do Anexo IV.

entre 20 e 30 anos, com o objetivo de avaliar se o aspecto finalizador dos enunciados se devia à entoação ou ao léxico. Os 36 pares de enunciados utilizados nesse teste foram deslexicalizados utilizando a técnica PURR (Sonntag & Portele, 1998). Nessa técnica o sinal acústico é deslexicalizado, mantendo intactos a amplitude e a informação de f_0 , portanto, a informação prosódica se manteve inalterada, fazendo com que fatores como informação lexical, significado, qualidade vocal e inteligibilidade da fala não influenciem o julgamento dos sujeitos.

Os testes que compõem o experimento foram realizados com o objetivo de determinar o grau de concordância dos sujeitos sobre as escolhas apresentadas. Os resultados do primeiro teste foram $k=0,12$ e $z=3,78$, o que indica uma concordância baixa e significativa entre os sujeitos no que diz respeito ao grau de ênfase. Os resultados do segundo teste, sobre a percepção do contorno HL como finalizador, não foram significativos, apresentando $k=0,04$, que indica concordância muito baixa, e $z=0,83$. No entanto, para o terceiro teste, com os *inputs* deslexicalizados, a concordância se mostrou melhor, mas ainda baixa, com $k=0,18$ e $z=4,18$.

Com esses contornos ocorre o mesmo caso que ocorre entre LH e >LH, ou seja, a diferença funcional dos contornos é baixa – talvez se aplicando a situações pragmaticamente específicas que não foram identificadas no experimento –, indicando um caso de compressão do movimento descendente ocasionada por uma restrição temporal.

i) > HL - late falling

O contorno >HL apresenta em termos gerais o mesmo movimento do contorno HL, porém em posição atrasada em relação ao movimento de f_0 na vogal tônica. Enquanto em HL a vogal tônica se alinha a posição baixa de f_0 , em >HL a vogal tônica se alinha ao movimento de descida de f_0 , como na figura 31.

Ocorrendo em menor escala que HL e em posição final de enunciados do PB e tem >HL o mesmo fenômeno que justifica o atraso em >LH, que é o desalinhamento entre

fonação e articulação. A figura 31 mostra a mudança no alinhamento entre o que seria um contorno HL e a realização do contorno >HL.

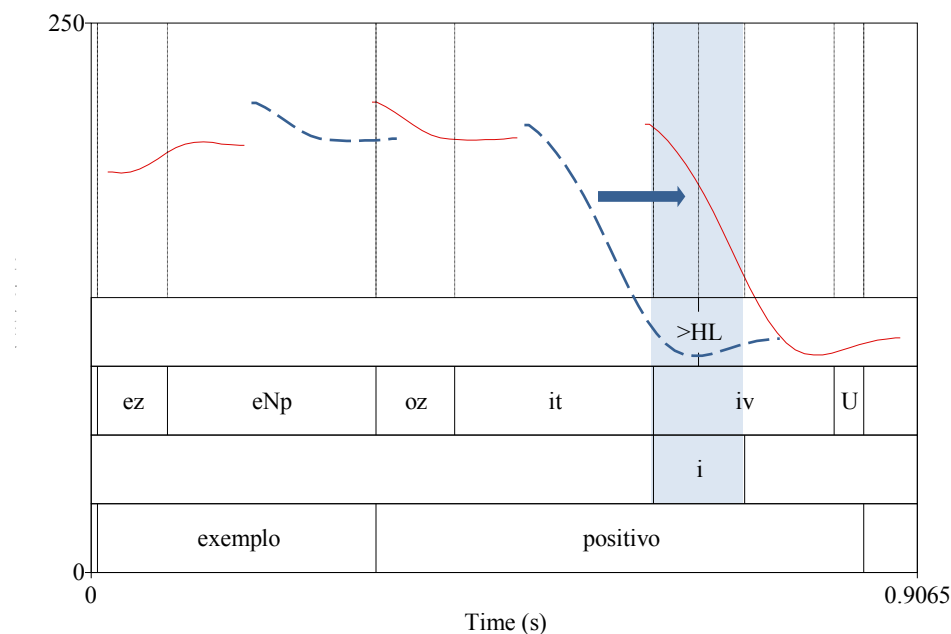


Figura 31: mudança no alinhamento entre o que seria um contorno HL (pontilhado) e a realização do contorno >HL alinhado à vogal tônica da palavra “positivo”.

Assim como ocorre entre LHL e HL, >HL apresenta a mesma diferença de movimento descendente em relação a LHL, que pode ser explicado por uma restrição temporal de >HL para atingir seu alvo. Esses contornos possuem um espaço de tempo para atingirem seu alvo, que é o alinhamento da posição baixa com a vogal tônica. A diminuição progressiva do espaço disponível para se atingir o alvo provoca uma compressão dos movimentos descendentes LHL, HL e >HL, resultando em mudanças nos contornos e no alinhamento. A figura 32 mostra o tempo disponível que cada contorno tem para atingir seu alvo – a vogal tônica – a partir de uma posição alta: LHL 63ms, HL 32ms e >HL 34ms. HL e >HL apresentam as mesmas medidas da posição alta até o alvo que é a posição baixa, pois tem o mesmo movimento, com diferença apenas no alinhamento. Na figura 33 é possível observar como esses contornos se organizam quanto ao alinhamento com a vogal tônica

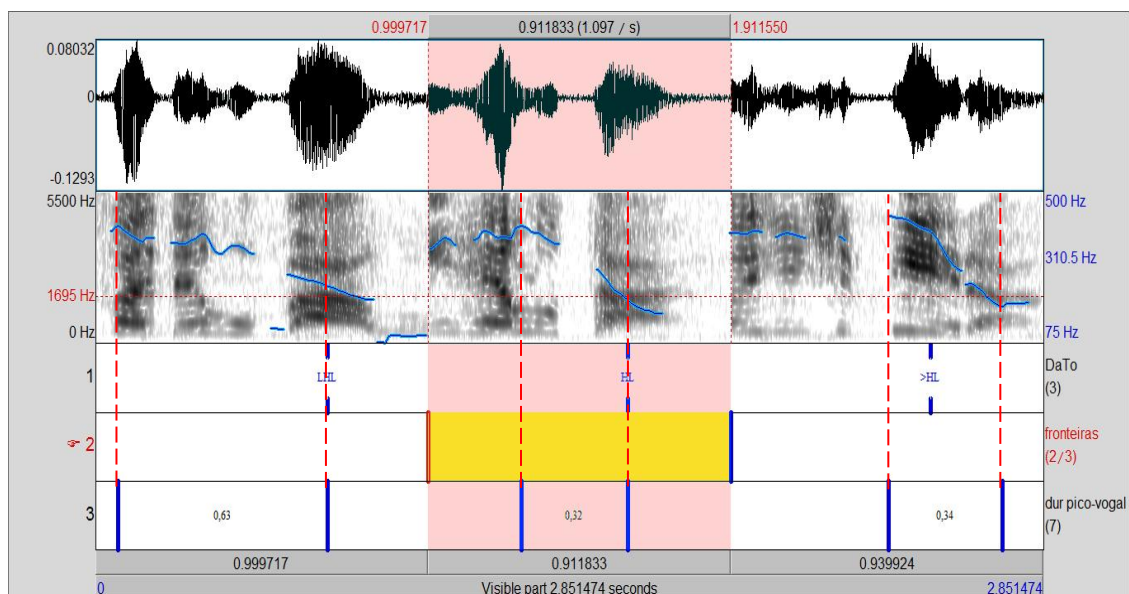


Figura 32: tempo que cada contorno tem disponível para atingir seu alvo a partir de uma posição alta.

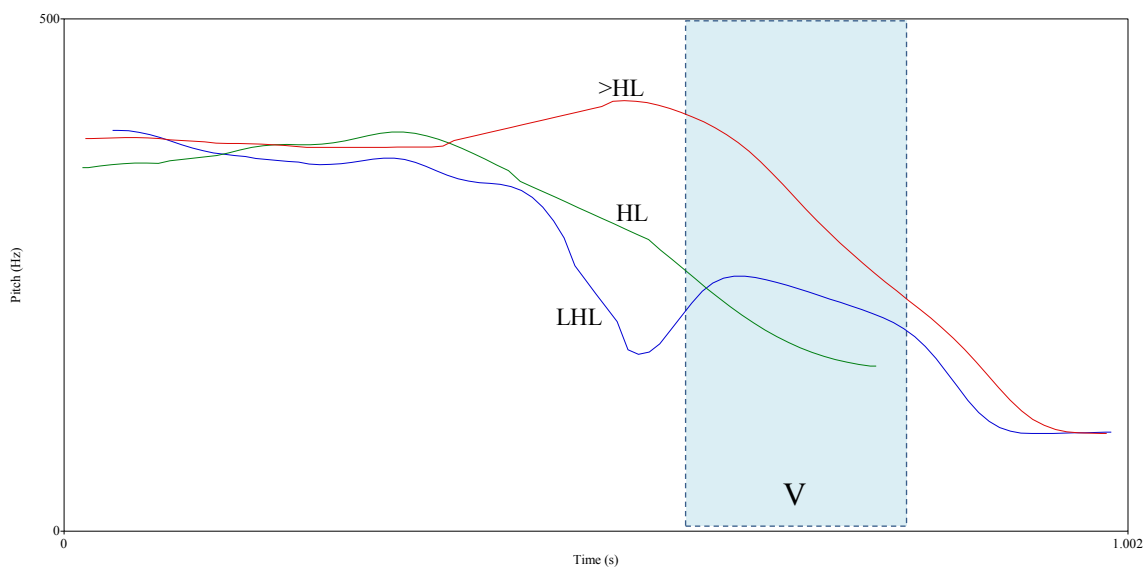


Figura 33: progressiva compressão dos contornos LHL (azul), HL (verde) e >HL (vermelho) sobrepostos e alinhados à vogal tônica

$v) > vHL - compressed falling$

O contorno vHL, assim com vLH é chamado de contorno comprimido, porém os fatores que determinam essa característica no contorno vHL são um pouco diferentes dos que atuam em vLH. vLH como dito anteriormente, ocorre entre dois picos de f_0 , fazendo com que haja uma restrição física e temporal para a realização de outro pico entre esses dois, por isso o alcance do alvo nesse ponto fica a cargo de um contorno comprimido. Com vHL, o que determina sua compressão são aspectos temporais, que levam à realização da descida de f_0 em um curto espaço, que não é suficiente para a organização completa da curva de HL, fazendo com que a descida ocorra bastante verticalizada e durante a realização da vogal tônica, sendo que em HL a descida ocorre durante a consoante da sílaba tônica. A figura 34 mostra dois exemplos do contorno vHL na repetição de uma mesma palavra.

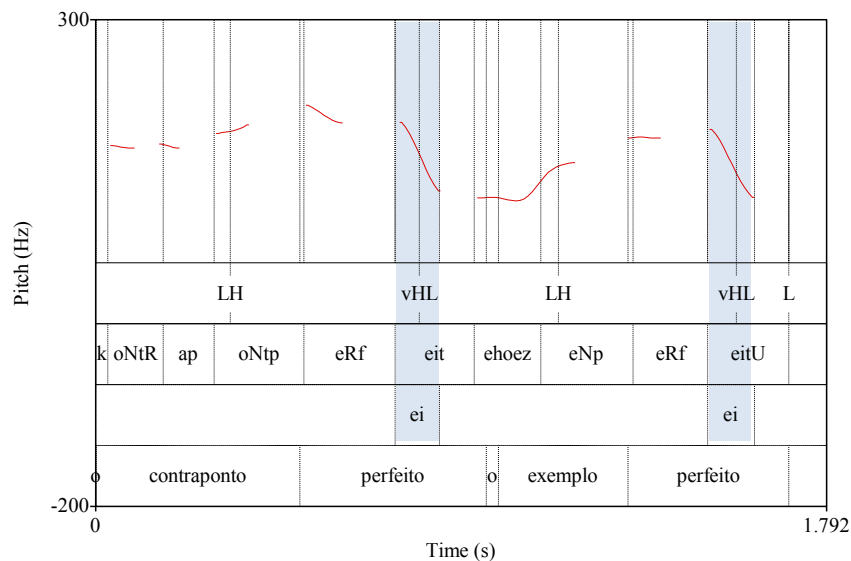


Figura 34: contorno vHL alinhado à descida de f_0 nas vogais tônicas da repetição da palavra “perfeito”.

O processo de compressão do contorno HL que dá origem ao contorno vHL é similar ao que ocorre entre LHL e HL, ou seja, a descida de f_0 para alcançar se alvo na posição baixa é determinada pelo espaço que o sistema dispõe para tanto, quanto menor esse espaço, mais os contornos tem que se ajustar ao espaço disponível. As consequências

dessa compressão são diferenças na função de cada contorno, pois como mostrou o experimento de percepção realizado, HL tem função diferente de LHL, e a distinção entre HL e vHL aparece no fato de vHL ocorrer em posição inicial de segmentos discursivos, o que indica que não possui a capacidade de finalização atribuída a HL.

1.3.1.3 Níveis de Fronteira

Os níveis de fronteira no sistema DaTo são representados apenas pelas letras H e L, e notam respectivamente níveis de fronteiras entoacionais altas e baixas. As fronteiras assinaladas no sistema DaTo são fronteiras finais de enunciados ou fronteiras que precedem pausas no interior dos enunciados. Fronteiras intermediárias como assinaladas no sistema ToBI são detectadas no sistema DaTo automaticamente pelo programa SG Detector (Barbosa, 2006), que faz a segmentação dos enunciados em grupos acentuais pela detecção de acentos frasais, que delimitam os grupos acentuais por terem duração maior na sílaba tônica²⁵. As figuras 35 e 36 ilustram a notação dos níveis de fronteira em fronteiras finais.

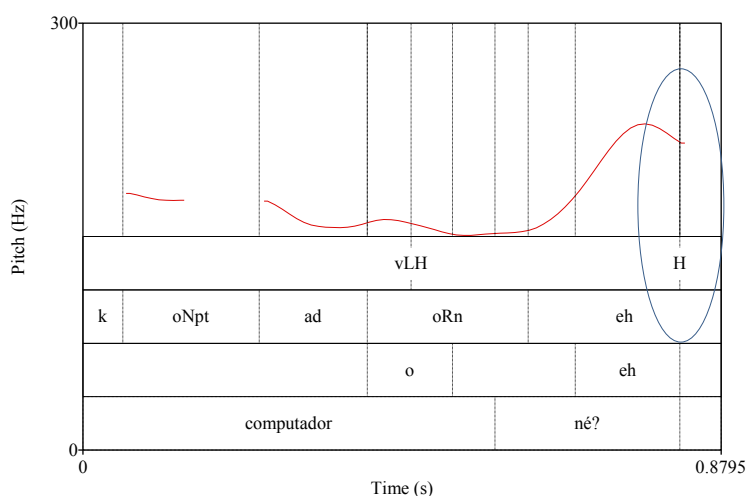


Figura 35: nível de fronteira H, após um mercador discursivo “né” na fronteira final do enunciado.

²⁵ Como se verá a seguir, o *SG Detector* será usado também para a marcação automática de fronteiras de grupos acentuais no sistema DaTo.

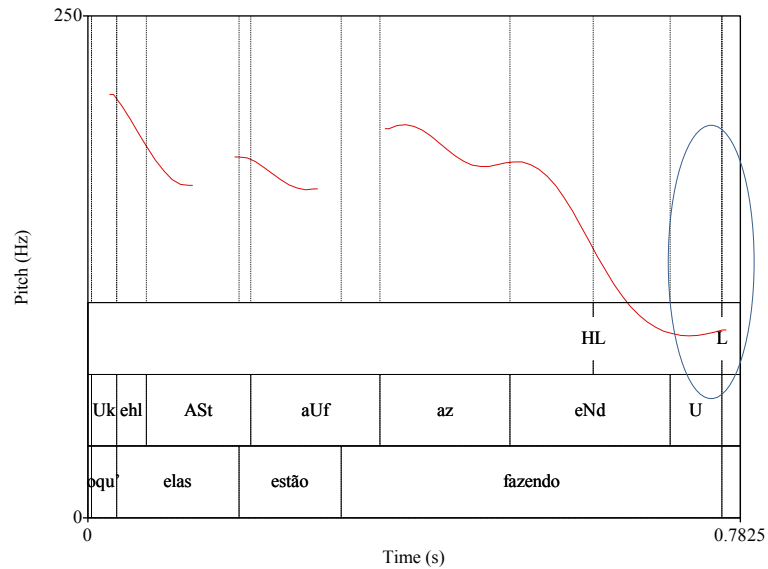


Figura 36: nível de fronteira L após contorno descendente HL.

A tabela 4 mostra um resumo dos contornos dinâmicos e dos níveis de fronteira do sistema DaTo.

Contornos Dinâmicos	
LH	<i>rising</i>
>LH	<i>late rising</i>
vLH	<i>compressed rising</i>
HLH	<i>falling-rising</i>
LHL	<i>rising-falling</i>
HL	<i>falling</i>
>HL	<i>late falling</i>
vHL	<i>compressed falling</i>
Níveis de Fronteira	
L	<i>low</i>
H	<i>high</i>

Tabela 4: contornos e níveis de fronteira do sistema DaTo.

1.3.2 Transcrição Ortográfica

A transcrição ortográfica que ocupa uma das camadas de notação do sistema é tão simplesmente a transcrição de como os enunciados foram realizados pelos participantes, com o intuito de preservar idiossincrasias da fala e do dialeto do falante. A transcrição ortográfica feita no programa Praat é alinhada item por item ao sinal de fala, e oferece a possibilidade de extração do texto no formato *.textgrid* para um arquivo *.txt* por meio de um *script* específico para essa finalidade (ver Anexo V). Dessa forma o conteúdo transcrito no sistema é idêntico ao da transcrição do corpus.

1.3.3 Notação Pragmática

A camada reservada para notação pragmática tem como função apontar fenômenos pragmáticos e comunicativos que possam facilitar a interpretação da ocorrência de fenômenos fonético-entoacionais, como realização de foco, interrupções (como pausas e hesitações) e alongamentos, por exemplo.

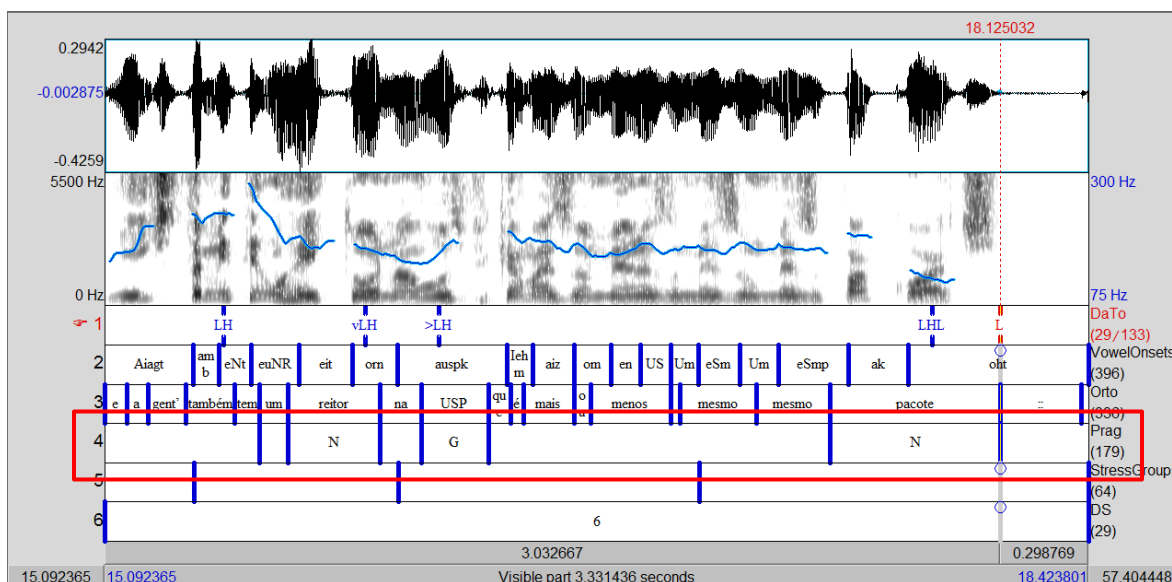


Figura 37: exemplo de notação utilizando a camada de notação pragmática do sistema DaTo.

A notação desses fenômenos não segue um formato específico, ficando a critério do pesquisador atribuir um rótulo para cada evento destacado. Em Lucente et al (2012) a ocorrência de informações novas e dadas foram anotadas nesta camada a partir da combinação dos seguintes rótulos, baseados na taxonomia proposta por Prince (1981) : N (informação nova no discurso, *new*) e G (informação já presente no discurso, ou *given*). Um exemplo dessa notação está na figura 37, com a camada destacada.

1.3.4 Unidades V-V

A camada da transcrição que apresenta a segmentação dos enunciados em unidades V- V compreende a divisão de todo enunciado em segmentos compreendidos entre os *onsets* de duas vogais. Essa segmentação é feita semi-automaticamente através do programa *Beat Extractor* (ver Barbosa, 2006, p. 451), que gera no programa Praat uma camada contendo a segmentação dessas unidades. Esse programa identifica os *beats*, pontos específicos associados ao aumento da amplitude na transição consoante – vogal (C-V). A detecção desses *beats* indica fronteiras próximas aos *onsets* das vogais e, portanto, auxilia na segmentação das unidades V-V.

Este tipo de segmentação no sistema DaTo, assim como proposto em Barbosa (2006) e Barbosa et al (2005), é usada para delimitar uma unidade cujas fronteiras são distintas das da sílaba fonológica e possui a vantagem de dinamizar a transcrição do sistema na medida em que a detecção de transições C-V por meio desse tipo de segmentação auxilia na observação de fenômenos de ordem física que podem estar relacionados à produção e percepção (Marcus, 1981; Dogil and Braun, 1988; Wong and Schreiner, 2003), e na determinação do ritmo (Barbosa, 2006) e da entoação.

De acordo com Barbosa (2006), as unidades V-V apresentam uma estabilidade duracional que oferece um efeito de compensação e uma extensão homogênea sobre a duração dos segmentos. Devido a essas propriedades, unidades V-V são usadas como blocos de construção para delimitação de grupos acentuais.

1.3.5 Grupos Acentuais

Os grupos acentuais (doravante SG do inglês *stress group*)²⁶ são unidades delimitadas por dois acentos frasais consecutivos. Os acentos frasais não se referem a aspectos sintáticos, mas sim a aspectos fonéticos, pois são definidos como proeminências no domínio da produção da fala, assinaladas por picos locais de duração ao longo dos enunciados (Barbosa, 2006, p.10). Por sua vez, os picos locais de duração que assinalam os acentos frasais são determinados pelos padrões de duração das unidades V-V. A relação entre unidades V-V, acentos frasais e grupos acentuais aparece ilustrada na figura 38.

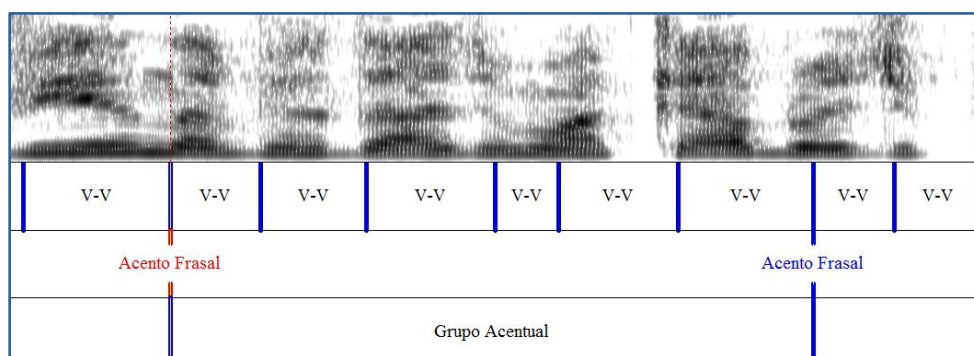


Figura 38: relação entre a segmentação em unidades V-V na primeira camada, acentos frasais na segunda camada e grupos acentuais (SGs) na terceira camada.

No sistema DaTo, os grupos acentuais são determinados automaticamente através do programa *SG Detector* (ver Barbosa, 2006, p. 459) que, assim como no caso das unidades V-V, gera uma camada com a segmentação dos grupos acentuais no programa Praat. O programa *SG Detector* detecta os acentos frasais pelo cálculo do desvio padrão das médias de duração das unidades V-V normalizadas pelo cálculo do *z-score* estendido (ver fórmula em Barbosa, 2006, p.169). Dessa forma o programa pode indicar onde estão os pontos de maior duração das vogais e, portanto, os acentos frasais.

A delimitação dos SGs através da detecção dos acentos frasais possibilita a localização e marcação automática das fronteiras entoacionais, as mesmas fronteiras que

²⁶ Aqui também será adotada a terminologia em inglês.

em ToBI são marcadas com os números 1, 2 e 3 baseando-se apenas na percepção (Beckman et al, 1994). A marcação dos SGs se mostrará bastante funcional no sistema DaTo, pois como se verá a seguir ela será responsável pela coordenação entre aspectos prosódicos e discursivos.

1.3.6 Segmentos Discursivos

A camada reservada para a segmentação discursiva tem como objetivo marcar as fronteiras entre cada enunciado de acordo com a divisão proposta no modelo G&S (Grosz & Sidner, 1986), que utiliza as expressões linguísticas como delimitadores entre cada enunciado que compõe um segmento discursivo (DS). Em outras palavras, é a transposição da segmentação discursiva efetuada no Corpus VoCE. para uma camada de notação.

1.4 Sumário

- o sistema DaTo contrapõe-se ao sistema ToBI, pois adota uma perspectiva dinâmica na notação entoacional, enquanto ToBI adota a teoria MA;
- o sistema DaTo não adota o conceito de *pitch accent* para notação dos contornos entoacionais, e sim o conceito de contorno dinâmico;
- o contorno dinâmico é definido como uma unidade tonal que contém elementos comunicativos expressos em uma trajetória da curva entoacional especificada por um alvo a ser atingido;
- a notação entoacional do sistema DaTo se divide entre contornos dinâmicos ascendentes, descendentes e níveis de fronteira;
- os contornos dinâmicos ascendentes e descendentes estabelecem seus padrões de movimento a partir de uma relação entre tempo e espaço para alcançar seu alvo;

- a notação pressupõe também transcrição ortográfica, segmentação dos dados em unidades V-V, marcação de grupos acentuais e segmentos discursivos.

2. INTERAÇÃO DISCURSO – RITMO - ENTOÇÃO

O sistema de notação DaTo, por não se limitar a fazer apenas a notação dos aspectos relacionados à curva entoacional, e acrescentar ao sistema ferramentas que dão conta de aspectos do ritmo da fala e do discurso, possibilitou a observação da existência de uma interação entre esses fatores de forma dinâmica.

Para justificar essa hipótese foi usado como modelo teórico o MDR (ver capítulo inicial), apresentado em Barbosa (2006), que apesar de se tratar de um modelo do ritmo, oferece um modelo plausível para a interação entre este, o discurso e a entoação, pois a análise dos dados de acordo com o sistema DaTo, assim como no MDR, indica um sistema que tem como mecanismo atrator os grupos acentuais. Além disso, Barbosa (2006) propõe a inclusão de um oscilador glotal em interação com os demais osciladores, proposta reformulada aqui.

2.1 Interações entre contorno entoacional, ritmo e discurso

Os dados do Corpus VoCE transcritos pelo sistema DaTo mostraram uma coincidência consistente entre fronteiras de grupos acentuais, contornos entoacionais e segmentos discursivos. Tais coincidências entre entoação, ritmo e discurso, ou entre aspectos prosódicos e discursivos, sugere a hipótese de que no MDR, paralelamente ao oscilador silábico, que determina a proeminência baseando-se apenas na duração dos segmentos, exista um oscilador glotal, que determina proeminências e fronteiras entoacionais de acordo com a frequência de vibração de pregas vocais na glote, sinalizados pelo movimento de subida ou descida de f_0 .

Num primeiro momento, foi observada a ocorrência de fronteiras de SG alinhadas proximamente aos contornos dinâmicos LH e >LH, e de nível de fronteira L. Os contornos LH e >LH são mais frequentes na marcação de foco estreito no PB, de acordo

com o sistema DaTo, o que indica que uma das estratégias para a marcação de proeminências é o alongamento na duração das vogais, indicado pelo alinhamento com o acento frasal.

Esse alinhamento não acontece de forma exata entre fronteiras de SG e contornos entoacionais. Pode-se dizer que se trata de um alinhamento aproximado, e isso se deve ao fato de que SGs e contornos entoacionais são marcados de forma distinta: os SGs são marcados com base na duração das unidades V-V que se utilizam da marcação dos *onsets* vocálicos para sua delimitação, enquanto os contornos entoacionais são anotados, por convenção, coincidindo com a posição medial das vogais tônicas. Dessa forma, o que consideramos alinhamento ocorre quando o contorno entoacional está marcado na vogal tônica da palavra em que ocorre o alongamento necessário para a detecção de uma fronteira de SG, como se pode observar na figura 39, que assinala o alinhamento entre os contornos >LH e LH e as fronteiras dos SGs.

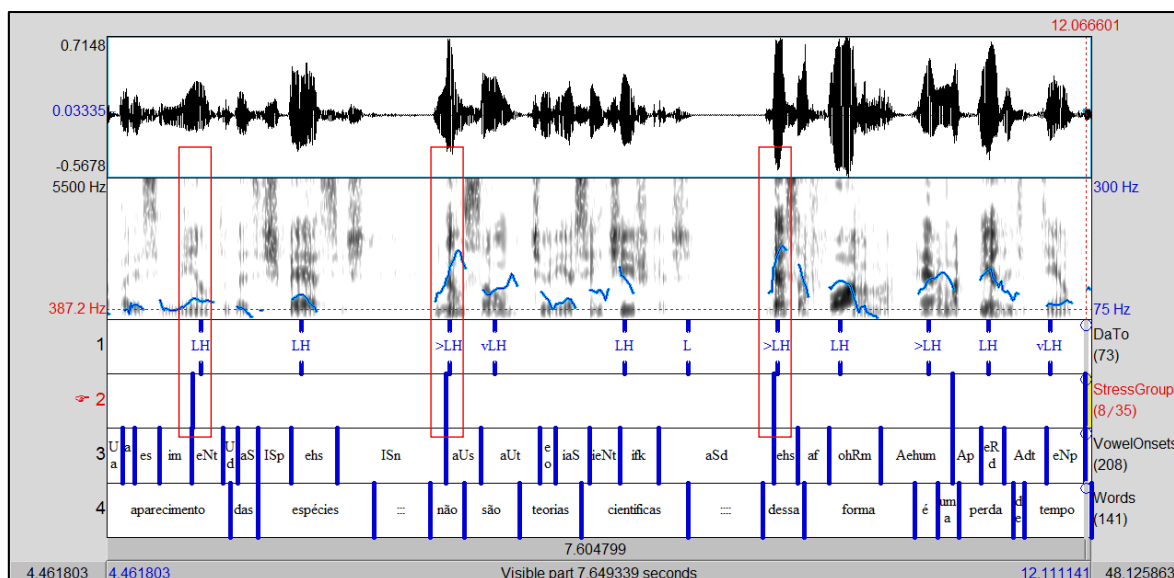


Figura 39: alinhamento entre fronteiras de grupos acentuais e contornos dinâmicos ascendentes

Em um segundo momento, após ter considerado pertinente para o estudo funcional dos aspectos entoacionais a análise da estrutura discursiva e levar a segmentação dos enunciados em unidades discursivas para as camadas de notação do sistema DaTo, foi

possível observar que existe um alinhamento aproximado também entre as fronteiras de DSs e SGs. Em geral a diferença no alinhamento que foi desconsiderada é do valor de duração de um segmento vocálico ou consonantal, não chegando a ser do valor de uma unidade V-V, pois, pelo mesmo motivo que determina a diferença no alinhamento entre SGs e contornos dinâmicos, a segmentação das unidades discursivas obedece fronteiras lexicais, enquanto nos SGs a segmentação é feita com base nas unidades V-V. Os casos em que o alinhamento entre SG e contornos dinâmicos se dá a uma distância acima do valor de uma unidade V-V foram considerados desalinhados.

Dos 315 segmentos discursivos analisados a partir do Corpus VoCE, foram encontrados 530 SGs, dos quais 56,7% (178) se alinham às fronteiras dos DS, e 43,3% (136) não se alinham. Os DSs têm duração média de 4,06s e contêm em média dois SGs, como mostram a tabela 5 e o gráfico de dispersão na figura 40. A tabela 4 mostra ainda outras medidas realizadas nesses segmentos. Esses dados incluem os SGs que não se alinham exatamente aos DSs, mas que se inserem nestes.

	Média	Mediana	Desvio Padrão
Duração dos DS	4,06 s	3,49	2,49s
SG por DS	2,74	2,0	1,77

Tabela 5: resumo das informações sobre a relação entre segmentos discursivos e grupos acentuais.

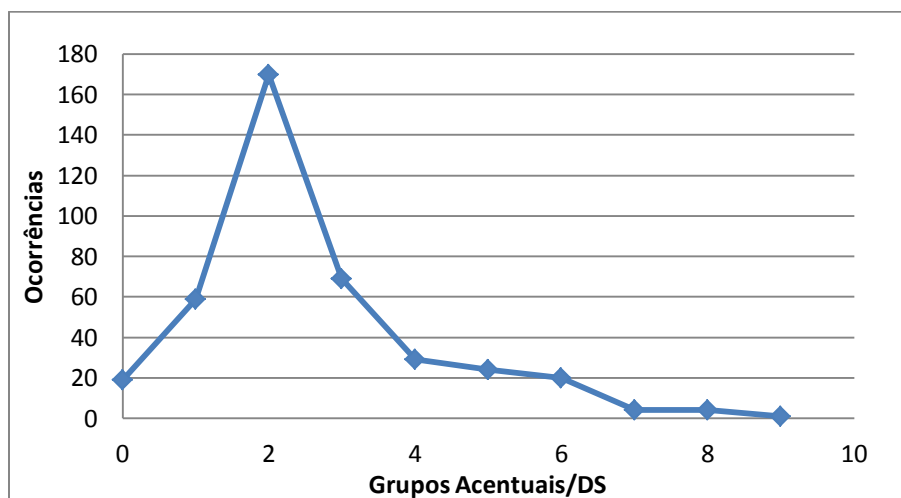


Figura 40: gráfico de distribuição e ocorrência de SG por DS; observe a concentração em torno de 2 SG/DS.

Além das interações entre ritmo e entoação e ritmo e discurso, os dados mostraram ainda a interação entre os três fenômenos. Como observado inicialmente, as fronteiras dos grupos acentuais coincidem em algum grau com a marcação dos contornos LH e >LH, assim como coincidem com a segmentação dos segmentos discursivos. Ao analisar o alinhamento entre essas três camadas foi possível observar que, no espaço de um segmento discursivo, a curva de f0 apresenta um padrão de movimento já descrito para enunciados declarativos em PB (Moraes, 1998; Lucente, 2008) que é a subida inicial de f0, marcada por LH ou >LH, seguida por um *resetting* de f0 até a fronteira final, marcada pelo nível L. O que foi observado nos dados de fala espontânea é que esse padrão se repete ao longo do discurso em alinhamento com os DSs, como mostra a figura 41.

Podemos observar que a cada início de DS ocorre um pico de f0, que contrasta com a fronteira baixa do segmento anterior, organizando o discurso em segmentos que atendem aos requisitos do modelo G&S, apresentados na seção anterior. Em números, a notação dos contornos iniciais de cada segmento dos 315 DSs é predominantemente marcada por LH e >LH, e a fronteira final pelo nível L, como se pode ver na tabela 6.

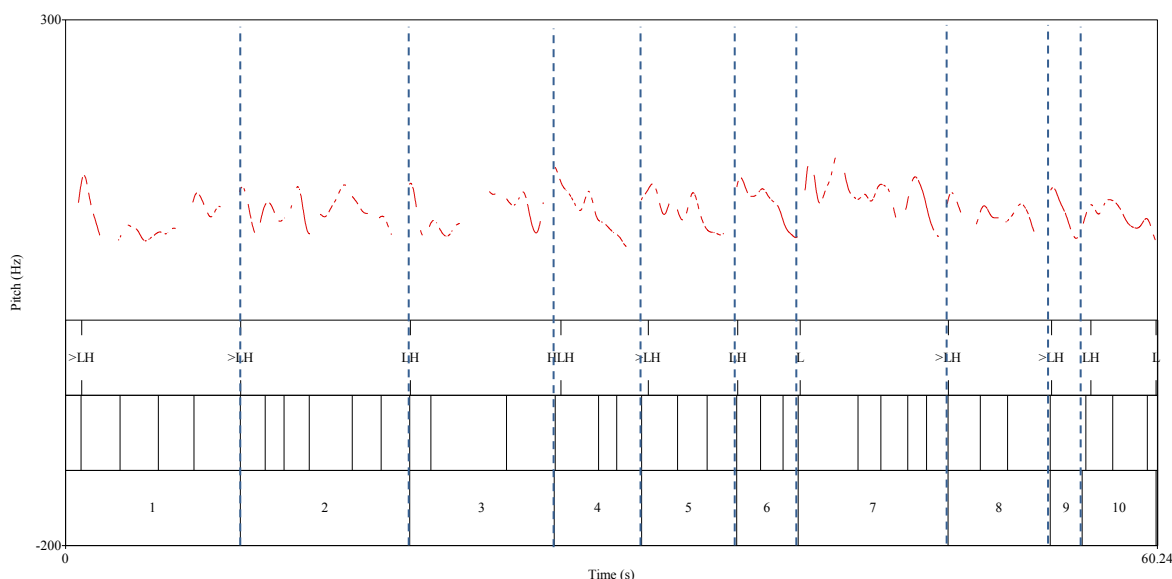


Figura 41: repetição de padrão de movimento ao longo da curva entoacional e alinhamento com os DSs; na primeira camada estão os contornos entoacionais, na segunda os SGs e na terceira os DSs.

		LH	>LH	vLH	HLH	LHL	>HL	vHL	HL	H	L	n ²⁷
Posição	Nº	149	106	1	6	1		2				50
Inicial	absolutos											
	%	47,30	33,65	0,31	1,90	0,31		0,63				15,87
Posição	Nº	24	18	2		32	3		11	49	153	18
Final	absolutos											
	%	7,61	5,71	0,63		10,15	0,95		3,49	15,55	48,57	5,71

Tabela 6: ocorrência de contornos dinâmicos na notação das posições iniciais e finais de enunciados que compõem os DSs.

Os DSs que não seguem esse padrão de f0 são encontrados em diferentes posições nos trechos de discurso, como se pode ver destacados na figuras 42 e 43 a seguir. Essa diferença de padrão entoacional pode ser atribuída à presença de enunciados interrogativos, pausas internas aos DSs, hesitações, correções – ou qualquer fenômeno que indique quebra, reorganização ou construção da sequência discursiva (ver segundo capítulo sobre Estrutura Discursiva) – presença de aspectos pragmáticos que indiquem suposição, presença de marcadores discursivos (*né?*). Na figura 42, que mostra a curva entoacional e sua suavização em 0,5 pontos, SGs e DSs do trecho de discurso *usp* (ver Anexo II), os DSs assinalados ilustram: no DS8 existem marcador discursivo e pressuposição sobre a presença de vários portões a partir de um alongamento e subida de f0 no final do enunciado (*porque né? a USP tem vários portões tem o portão principal :::*); no DS16 ocorre pausa e presença de marcador discursivo final (*e o pessoal fala pô e ó :::: governo deixou explorar isso aqui né? :::*), nesse segmento é interessante observar que ocorre um *reset* de f0 após a pausa intermediária; no DS18 ocorrem pausa, hesitação e ironia sobre uma suposta solução para os atrasos devido ao fato dele não poder atravessar a USP de carro (*eu fiquei pensando né? cara eh :: os caras podiam dar pelo menos um atestado pra eu levar pro empregador*).

O mesmo ocorre na figura 43 com o trecho de discurso *trama*. Os DSs 3, 7 e 13 assinalados contêm, respectivamente: hesitação, pausa, enumeração (*porque eu lembro que ah ah na época que eu retornei ao Brasil ::: o que realmente interessava pras grandes*

²⁷ N se refere à ocorrência de outros marcadores que não os contornos dinâmicos, como hesitações, pausas, etc.

gravadoras ou era o sertanejo :::: ou o axé : ou o pagode :: qualquer uma :); marcador discursivo, pressuposição (então :::: eh a Trama junto com outras gravadoras independentes começaram apostar em outras coisas); pausa, hesitação (eh mas eu acho que a importância ::: da n...da nossa geração e da Trama e dessa mentalidade do do mercado dependente musical :: foi justamente abrir um espaço pra coisas que não tavam rolando).

A correlação cruzada²⁸ das notações dos DSs, feita no *software* livre Weka (3.6.5) mostra em números absolutos, que a combinação dos contornos dinâmicos LH e >LH em posição inicial e de L em fronteiras de enunciados também é predominante (36%), como mostra a tabela 7. A aplicação do teste chi-quadrado para variáveis independentes ($\chi^2 = 78.22$; $df = 54$; $p = 0.01724$) mostrou resultado significativo para variáveis dependentes.

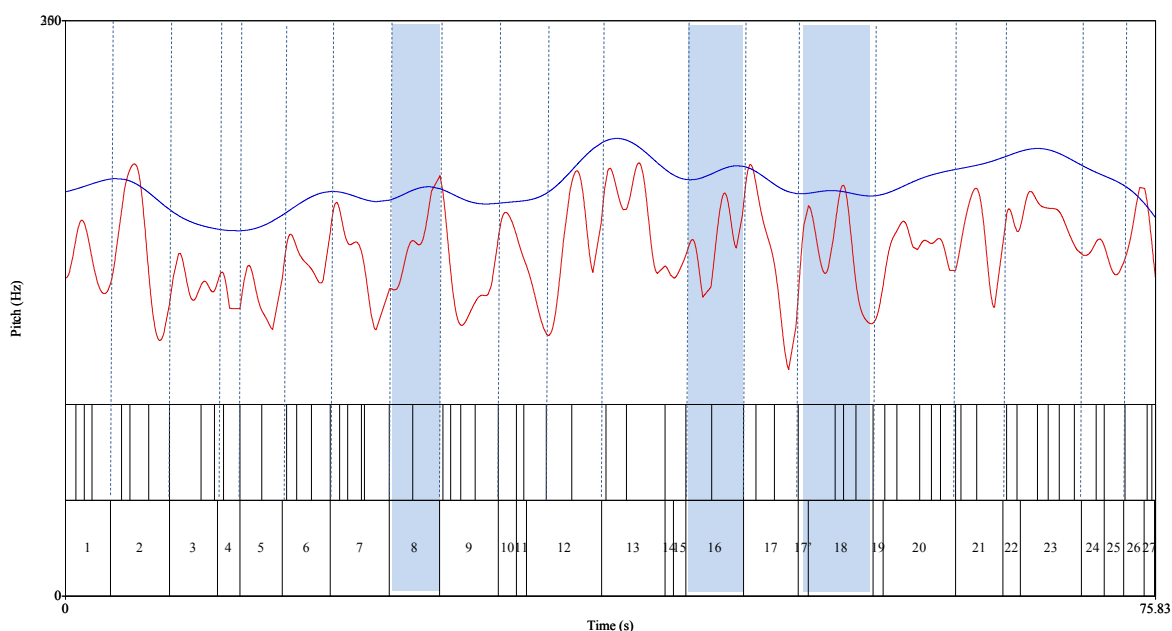


Figura 42: exceções no padrão de movimento da curva entoacional (f0 em vermelho e suavização de f0 em 0.5 pontos em azul) e alinhamento DSs no trecho *USP*

²⁸ A função correlação cruzada empregada aos dados do corpus remete o quanto os dois processos, até então independentes, estão correlacionados entre si. Essa função é bastante empregada em séries temporais na avaliação da correlação entre eventos ao longo do tempo.

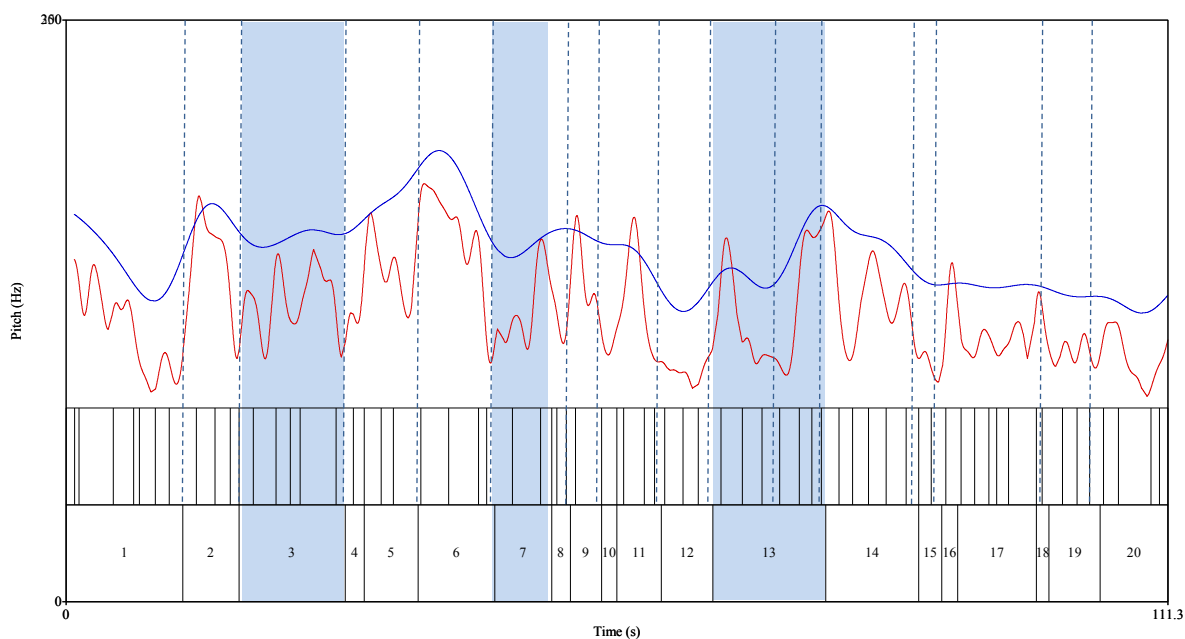


Figura 43: exceções no padrão de movimento da curva entoacional (f_0 em vermelho; f_0 suavizado em azul) e alinhamento DSs no trecho *trama*

Contorno Inicial	LH	>LH	n	HLH	vHL	vLH	LHL
Contorno Final							
L	80.0	62.0	10.0	4.0	2.0	2.0	1.0
>HL	2.0	3.0	1.0	1.0	1.0	1.0	1.0
>LH	9.0	6.0	6.0	1.0	1.0	1.0	1.0
H	26.0	15.0	8.0	3.0	2.0	1.0	1.0
HL	6.0	4.0	5.0	1.0	1.0	1.0	1.0
HLH	1.0	2.0	3.0	1.0	1.0	1.0	1.0
LH	13.0	8.0	5.0	2.0	1.0	1.0	1.0
LHL	15.0	11.0	9.0	1.0	1.0	1.0	1.0
n	5.0	4.0	11.0	1.0	1.0	1.0	2.0
vLH	2.0	1.0	2.0	1.0	1.0	1.0	1.0

Tabela 7: correlação cruzada das ocorrências de combinação de contornos iniciais e finais para cada um dos 315 enunciados que compõem os DSs.

Esses números sugerem que a repetição do padrão do movimento da curva entoacional, iniciado por um pico e finalizado em posição baixa, alinhada aos segmentos

discursivos, que por sua vez se alinham às fronteiras de grupos acentuais em grande parte dos casos, mantém uma relação consistente entre os três níveis de análise. De acordo com o MDR, a observação de como se estrutura essa relação sugere que o oscilador silábico pode estar em acoplamento com um mecanismo que gera padrões glotais, e consequentemente sugere também a hipótese de que no PB os falantes organizam a fala, ou o discurso, em ciclos temporais ancorados em padrões rítmico-entoacionais.

Em suma: os dados de segmentação discursiva, segmentação em grupos entoacionais e notação entoacional revelam uma estrutura dinâmica que concentra em grande maioria dois SG a cada DS com duração aproximada de 4s. As fronteiras dos DSs coincidem em 56% com fronteiras de SGs, ou seja, com os acentos frasais. Os acentos frasais, determinados pelo aumento na duração dos segmentos vocálicos também se alinham aos contornos entoacionais LH e >LH em posição inicial dos DSs e aos contornos L e LHL em posição final dos DSs. A correlação cruzada entre os contornos entoacionais mostra forte correlação entre LH – L, entre >LH – L e entre LH – LHL, que indica um segmento iniciado com contorno ascendente e finalizado com um *resetting* de f0, formando um padrão que se repete ao longo do discurso²⁹.

Esse alinhamento entre as estruturas discursiva, rítmica e entoacional é possível por uma coordenação de fatores que convergem para a determinação dos SGs como atratores de um sistema produtor da fala e da entoação.

2.2 Nova Hipótese para o Modelo Dinâmico do Ritmo

Da hipótese da organização do discurso em ciclos temporais ancorados em padrões rítmico-entoacionais, e da existência de um oscilador glotal em sincronia com o oscilador silábico no MDR, surge a proposta de acréscimo deste oscilador no MDR, já apresentado por Barbosa (2006).

²⁹ A hierarquia na segmentação discursiva não foi explorada nesta análise, a proposta futura é analisar como esses níveis hierárquicos se organizam entoacionalmente usando o mesmo tipo de análise.

Em nossa proposta de reformulação, o oscilador glotal se encaixa no MDR paralelamente ao oscilador acentual, interagindo com informações contidas em níveis linguísticos elevados – onde se inserem as informações sintática, semântica e pragmática do discurso – e acoplado ao oscilador silábico por uma força de acoplamento (x). O oscilador silábico mantém sua relação com o oscilador acentual e com a pauta gestual da mesma forma. O que é proposto nessa inserção é que o acoplamento entre os osciladores glotal e silábico possa determinar o alinhamento entre informações discursivas contidas nos padrões entoacionais e grupos acentuais.

Nessa nova proposta o oscilador glotal mantém ainda uma relação de acoplamento (z) com o oscilador acentual. Essa relação determina o alinhamento entre o alvo dos contornos entoacionais e as sílabas tônicas. Ou seja, o oscilador glotal se relaciona com os osciladores silábico e acentual, num sincronismo que une duração, tonicidade e movimento de glote na determinação da ênfase e dos padrões entoacionais do discurso. O diagrama da figura 44 ilustra a nova proposta para o MDR.

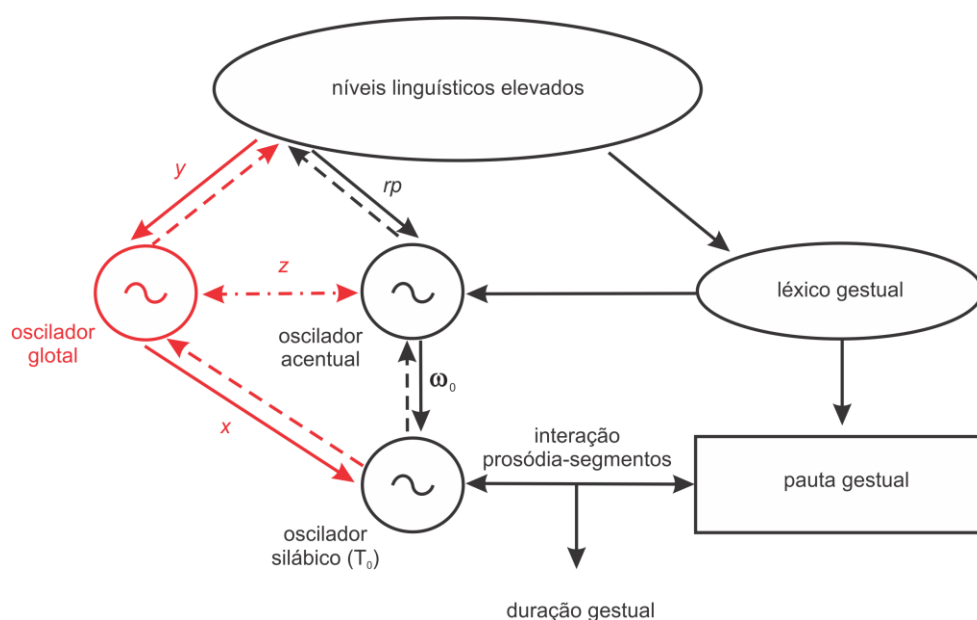


Figura 44: diagrama do Modelo Dinâmico do Ritmo da Fala

O sistema DaTo vem usando a segmentação dos enunciados em unidades V-V em sua notação desde sua implementação, e se valendo dessa segmentação para observações de como nos momentos de transição C-V se alinham a picos de velocidade na subida e/ou descida de f_0 , alinhamento de picos de f_0 com vogais tônicas, transições nos movimentos de f_0 encoradas nas transições C-V e V-C, etc.

A observação de como se estrutura essa relação entre os padrões de duração das unidades V-V, como resposta à ação do oscilador silábico e a organização entoacional, que aponta para a existência de um oscilador glotal, sugere a hipótese de que, no PB, os falantes organizam a fala em ciclos ou grupos que seguem padrões entoacionais. Esse tipo de observação sobre o comportamento da curva entoacional do PB auxilia na delimitação de um programa de investigação científica dinamicista, pois corrobora a hipótese da ação de diferentes sistemas num único sistema que os contém.

De acordo com o que propõem as teorias que trabalham com sistemas dinâmicos, a entoação pode ser vista como resultante da interação entre diversos elementos em um sistema dinâmico-funcional, como propõe a fonologia articulatória (Browman e Goldstein, 1986; Saltzman, 1995), e o estudo da produção do ritmo da fala (Barbosa, 2006). Essa incursão no campo da fonologia articulatória, ao ser adotado o modelo de Barbosa (2006), preenche uma lacuna que o modelo PENTA, proposto por Xu (2005) não preenchia e que se faz necessária no tipo de abordagem que o projeto faz dos aspectos entoacionais, que é a necessidade de representação dos fenômenos entoacionais. O modelo PENTA não pressupõe representação fonológica para os eventos entoacionais, considerando estes estritamente fonéticos. A perspectiva adotada aqui, apesar de tratar a entoação sob um ponto de vista fonético, pressupõe ainda a representação dos contornos entoacionais por meio do sistema DaTo, que é ferramenta central na pesquisa a que se propôs esta tese.

O MDR, apesar de não conter uma representação fonológica nos moldes de um sistema de notação, não torna excludente o papel da fonologia na representação prosódica, muito embora considere a relevância de um alinhamento gradiente entre eventos singulares dos contornos representados com o material linguístico. No programa em que o MDR se insere há espaço para abstração (fonologia), pois essa informação abstrata é comensurável como os sistemas físicos/motores que realizam esses padrões abstratos.

Do ponto de vista da existência do oscilador glotal no MDR, o sistema DaTo tem a função de reconciliar a geração do sinal glotal com a representação fonológica deste.

2.3 Sumário

- a notação DaTo mostra que as fronteiras dos DSs se alinham com fronteiras de SGs e que os SGs se inserem nos DSs numa proporção de dois SGs a cada DS;
- os DSs encerram um padrão entoacional formado por contornos iniciais LH e >LH em forte correlação com o nível de fronteira L e com o contorno LHL;
- o alinhamento entre estruturas discursiva (DSs), rítmica (SGs) e entoacional (contornos dinâmicos) indica um sistema dinâmico de produção da fala e da entoação que tem o SG como atrator;
- da sincronia entre ritmo e entoação surge a proposta de inserção de um oscilador glotal no MDF;
- o sinal glotal no MDF tem sua representação fonológica no sistema DaTo.

IV CONCLUSÃO

CONCLUSÃO

Esta tese pode ser dividida em duas partes, uma dedicada ao discurso e outra dedicada à entoação. Essas partes se complementam na medida em que a estrutura de análise discursiva e segmentação das unidades que compõem o discurso passam a ser analisadas por meio de uma ferramenta comum à análise entoacional: o sistema DaTo.

O fato de uma estrutura discursiva poder ser analisada de tal forma se deve à abordagem dinâmica adotada para o estudo e análise da entoação. A partir do momento em que se adota um programa de pesquisa que abandona a dicotomia entre mente e corpo é possível unir em um único sistema cognitivo-motor a elaboração do discurso e a produção da fala.

O sistema DaTo, em sua notação, agrega esse fator dinâmico na medida em que possibilita a observação da emergência de padrões entoacionais relacionados aos segmentos discursivos. Paralelamente a essa relação do discurso com a entoação, o ritmo age como um atrator e coordena as três estruturas: discursiva, rítmica e entoacional.

Se o sistema DaTo oferece condições para a observação da coordenação dessas estruturas e a notação dos elementos que fazem transparecer a economia dessa coordenação, o Modelo Dinâmico do Ritmo, proposto por Barbosa (2006), com a proposta de inclusão de um oscilador glotal, assume a tentativa de explicar o funcionamento dessa organização em um sistema dinâmico.

A hipótese da existência de um oscilador glotal, que funciona recebendo informações pragmáticas e discursivas de um nível superior – sistema cognitivo – e sua relação com os osciladores acentual e silábico – sistema motor –, pode explicar o fato de a informação contida no discurso ocorrer em um segmento específico, com acentuação e duração apropriadas para que tenha seu significado compreendido no discurso.

A contribuição dessa tese é a de oferecer uma visão dinâmica do funcionamento da entoação e do discurso, algo que se aproxima mais da interação entre mecanismos físicos e funcionais do que de uma subjetividade do sujeito. O estudo da entoação da fala mostra que esta agrega funções necessárias para o entendimento entre os falantes e pode ser

revelador de outros aspectos dinâmicos presentes no funcionamento da fala que ainda estão por ser investigados.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abner, N. 2009. Phrasing and Prominence: Extracting Prosodic Information from the BU Radio Corpus. Ms. UCLA.
- Atterer, M. and Ladd, D. R. (2004) On the phonetics and phonology of "segmental anchoring" of f_0 : evidence from German. *Journal of Phonetics* 32, 177-197.
- Barbosa, P. A. (2006). *Incursões em torno do ritmo da fala*. Campinas: Pontes.
- _____. (2007) From syntax to acoustic duration: a dynamical model of speech rhythm production. *Speech Communication*. 49 (1-2), 725 - 742.
- _____. (2008) Prominence- and boundary-related acoustic correlations in Brazilian Portuguese read and spontaneous speech. *Proceedings of the Speech Prosody 2008 Conference*. Campinas.
- _____. (2009). *Prosódia*. Enciclopédia Virtual de Psicolinguística. Acessível em: <http://wikipsicolinguistica.org.br>
- _____. (2010) Automatic duration-related salience detection in Brazilian Portuguese read and spontaneous speech In *Proceedings of the Speech Prosody 2010*. Chicago.
- Barbosa, P. A., Arantes, P., Meireles, A. R., Vieira, J. M. (2005). Abstractness in Speech-Metronome Synchronisation: P-Centres as Cyclic Attractors. *Proceedings of the Ninth European Conference on Speech Communication and Technology (Interspeech 2005)* Lisbon, Portugal (1441-1444).
- Baumann, Stefan & Arndt, Riestter (2010). Annotating Information Status in Spontaneous Speech. *Proceedings Speech Prosody 2010*, p. 1-4.
- Baumann, S. (2006). "The Intonation of Givenness - Evidence from German". *Linguistische Arbeiten*, Tübingen: Niemeyer (PhD thesis, Saarland University).
- Beckman, M. E. (1996). A typology of spontaneous speech. In Y. Sagisaka, W. N. Campbell & N. Higuchi, eds., *Computing Prosody*, pp. 7-26 . New York: Springer-Verlag.
- Beckman, M. E., Hirschberg, J., Pitrelli, John F., (1994). *Evaluation of Prosodic Transcription Labeling Reliability in the ToBI Framework*. (Disponível em http://www.ling.ohio-state.edu/~tobi/ame_tobi).
- Biber, D., Conrad, S. (2009) *Register, genre and style*. Cambridge University Press.
- Boersma, P., Weenink, D. (2009): Praat: doing phonetics by computer (Version 5.1.05) [Computer program]. Retrieved May 1, 2009, from <http://www.praat.org/>
- Borges Neto, J (2004) *Ensaio de filosofia da linguística*. São Paulo, SP: Parábola.
- Botinis, A., Granström, B., Möbius, G., (2001) Developments and paradigms in intonation research. *Speech Communication* 33(4): 263-296.

- Brown, G. (1983) "Prosodic structure and the given_new distinction". In Ladd, D. R. & Cutler, A. (eds.) *Prosody: Models and Measurements*, Springer Verlag, Berlin, p.67-78.
- Browman, C.P. & Goldstein, L. (1986) Towards an articulatory phonology. In C. Ewan and J. Anderson (eds.) *Phonology Yearbook 3*. Cambridge: Cambridge University Press, p. 219-252.
- Butterworth, B. (1975) Hesitation and semantic planning in speech. In *Journal of Psycholinguistic Research* (4); p.143-178.
- Campbell, W.N. (1996) Synthesizing Spontaneous Speech. In Sagisaka, Y., Campbell, N., Higuchi, N. (eds) *Computing Prosody*. Computational Models for Processing Spontaneous Speech. New York: Springer-Verlag.
- Chafe, W. (1976) "Givenness contrastiveness de_niteness subjects topics and point of view". In Li, C., (ed.) *Subject and Topic*, Academic Press New York, p. 25-55.
- Chafe, Wallace L. (1979) The Flow of Thought and the Flow of Language. In Givon. T., Ed., *Syntax and Semantics, Vol. 12, Discourse and Syntax*. Academic Press, New York, New York: 159- 182.
- Chafe, W.L. (1980) The Deployment of Consciousness in the Production of a Narrative. In Chafe, W.L., Ed., *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production. Vol. 3. Advances in Discourse Processes*. Ablex Publishing Corp, Norwood, New Jersey: 9-50.
- Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. O. Freedle (Ed.), *Discourse production and comprehension* (pp. 1-40). Hillsdale, NJ: Erlbaum.
- Cohen, P. R., and Levesque, H. J. (1980). Speech Acts and the Recognition of Shared Plans. In *Proceedings of the 3rd Conference of the Canadian Society for Computational Studies of Intelligence*, Victoria, B.C., p. 263-271.
- Côrtes, P. O., Mittmann, M. M., Caetano, R. V. O., Mello, H. R.; Raso, T. (2011) A convergência entre anotadores na segmentação prosódica do corpus C-ORAL-BRASIL. In: *Anais do III Colóquio Brasileiro de Prosódia da Fala*. Belo Horizonte, p. 1-7.
- Dogil, G., Braun, G. (1988) The PIVOT model of speech parsing.. Vienna, Áustria: Verlag.
- Escudero, D., Aguilar, L., del Mar Vanrell, M., Prieto, P. (2012) Analysis of inter-transcriber consistency in the Cat_ToBI prosodic labeling system. *Speech Communication* 54 (2012), p. 566–582.

- Fávero, L. L., Andrade, M. L. C. V. O., Aquino, Z. G. O. (2006) “Correção”. In Jubran, C. C. A. S. & Koch, I. G. V. (Org.) (2006). *Gramática do português culto falado no Brasil*. Volume 1: Construção do texto falado. Campinas: Editora da Unicamp.
- Fernandes, F. (2007) Ordem, focalização e preenchimento em português: sintaxe e prosódia. . Ph.D. Thesis. Unicamp, Campinas.
- Fujimura, O. (2000) The C/D model and prosodic control of articulatory behavior. *Phonetica* 57, p.128-138.
- Gravano, A. and Hirschberg, J. (2006) "Effect of Genre, Speaker, and Word Class on the Realization of Given and New Information," Proceedings of Interspeech 2006, Pittsburgh.
- Grice, M., Ladd, D. R., Arvaniti, A. (2000) On the place of phrase accents in intonational phonology. *Phonology* 17: p. 143-185.
- Grosz, B.J., and Sidner, C.L., (1986) "Attention, Intentions, and the Structure of Discourse", *Computational Linguistics*, p. 12:3.
- Hasegawa-Johnson, Mark, Chen, Ken, Cole, Jennifer, Borys, Sarah, Kim, Sung-Suk, Cohen, Aaron, Zhang, Tong, Choi, Jeung-Yoon, Kim, Heejin, Yoon, Tae-Jin, & Chavarria, Sandra. (2005) Simultaneous Recognition of Words and Prosody in the Boston University Radio Speech Corpus. *Speech Communication* 46, p. 418-439.
- Hirschberg, J., and Litman, D. (1993) Empirical Studies on the Disambiguation of Cue Phrases, *Computation Linguistics*, 19-3, p. 501–530.
- Hirst, D. (2005). Form and function in the representation of speech prosody. *Speech Communication* 46(3-4): p. 334-347.
- Jubran, C. C. A. S. & Koch, I. G. V. (Org.). (2006). *Gramática do português culto falado no Brasil*. Volume 1: Construção do texto falado. Campinas: Editora da Unicamp.
- Jubran, C. C. A. S. (2006) “Parentetização”. In Jubran, C. C. A. S. & Koch, I. G. V. (Org.) (2006). *Gramática do português culto falado no Brasil*. Volume 1: Construção do texto falado. Campinas: Editora da Unicamp.
- Kelso, S. (1984). Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology: Regulatory, Integrative and Comparative* 15.
- _____ (1995). *Dynamic Patterns*. Cambridge: MIT Press.
- Koch, I. G. V. (2006) “Tematização e rematização”. In Jubran, C. C. A. S. & Koch, I. G. V. (Org.) (2006). *Gramática do português culto falado no Brasil*. Volume 1: Construção do texto falado. Campinas: Editora da Unicamp.

- Kohler, K. J. (1996) Modelling Prosody in Spontaneous Speech. In Sagisaka, Y., Campbell, N., Higuchi, N. (eds) *Computing Prosody. Computational Models for Processing Spontaneous Speech*. New York: Springer-Verlag.
- Kohler, K. J. (2005) Timing and Communicative Functions of Pitch Contours. *Phonetica* 62, p. 88-105.
- Kugler, P. N., Turvey, M. T. (1987). Information, natural law, and the self-assembly of rhythmic movement. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Kuhn, T. S. (1987) *A estrutura das revoluções científicas*. São Paulo: Perspectiva.
- Ladd, D. R. (1996a) *Intonational Phonology*. Cambridge: Cambridge University Press.
- _____. (1996b) Introction. In Y. Sagisaka & al. (Eds). *Computing Prosody*, New-York: Springer: p. 67-80.
- Lakatos, I. (1978) *Falsificação e Metodologia dos Programas de Investigação Científica*. Lisboa: Edições 70.
- Lehiste, I., (1970) *Suprasegmentals*. Cambridge: MIT Press.
- Lieberman, M. (1975) *The intonational system of English*. Ph.D. Thesis; MIT.
- Lucente, L., Hirschberg, J., Barbosa, P. A. (2012) “The role of discourse structures and intonational features on determining the information status prominence” (*to appear*).
- Lucente, L. (2008) *DaTo: Um sistema de notação entoacional do português brasileiro baseado em princípios dinâmicos. Ênfase no foco e na fala espontânea*. Dissertação de Mestrado. Unicamp.
- Lucente, L., Barbosa, P. A. (2010) The role of alignment and height in the perception of LH contours. *Proceedings of Fifth Conference on Speech Prosody*. Chicago.
- _____. (2008) Narrow focus in Brazilian Portuguese: spatial and temporal constraints. *Proceedings of Fourth Conference on Speech Prosody*. Campinas.
- _____. (2007) Notação Entoacional do Português Brasileiro em Corpora de fala Semi-Espontânea e Espontânea. *Revista Intercâmbio* 16.
- Marcuschi, L. A., (2006) “Referenciação”. In Jubran, C. C. A. S. & Koch, I. G. V. (Org.) (2006). *Gramática do português culto falado no Brasil*. Volume 1: Construção do texto falado. Campinas: Editora da Unicamp.
- Marcuschi, L. A., Koch, I. G. V. (2006) “Hesitação”. In Jubran, C. C. A. S. & Koch, I. G. V. (Org.) (2006). *Gramática do português culto falado no Brasil*. Volume 1: Construção do texto falado. Campinas: Editora da Unicamp.

- Merlo, S. (2012). Dinâmica temporal de pausas e hesitações na fala semi-espontânea. Tese de Doutorado. Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, SP.
- Moraes, J. A. (1998) “Intonation in Brazilian Portuguese”. In Hirst, D., Di Cristo, A. (eds.) *Intonational Systems: a Survey of Twenty Languages*. Cambridge. MIT Press.
- Nakatani, C. H. (1996). Integrating prosodic and discourse modelling. In Y. Sagisaka & al. (Eds). *Computing Prosody*, New-York: Springer: p. 67-80.
- Nakatani, C. H., Hirschberg, J. (1994) A Corpus-based study of repair cues in spontaneous speech, *Journal of the Acoustical Society of America*, 95-3: p. 1603–1616.
- Ostendorf, M., Price, P., and Shattuck-Hufnagel, S. (1996) “Boston University Radio Speech Corpus”, Linguistic Data Consortium, Philadelphia.
- Pierrehumbert, J. (1980) The Phonology and Phonetics of English Intonation. Ph.D thesis, MIT.
- Pires de Oliveira, R. (2010) A linguística sem Chomsky e o método negativo. *ReVEL*, vol. 8, n. 14.
- Polanyi, L. and Scha, R.J.H. (1986) “Discourse Syntax and Semantics”. In Polanyi, L., Ed., *The Structure of Discourse*. Ablex Publishing Co., Norwood, New Jersey.
- Popper, K. R. (1996) *A lógica da pesquisa científica*. Imprensa São Paulo : Cultrix.
- Prince, E. F., (1981) “Toward a taxonomy of given-new information”, In Peter Cole (ed.), *Radical Pragmatics*, The Academic Press, New York, p. 223-255
- Raso, T., Mello, H., (2012) C-ORAL-BRASIL I: Corpus de referência do português brasileiro falado informal. Belo Horizonte, UFMG-CNPq.
- Rietveld, A. C. M., Gussenhoven, C. (1988) “On the relation between pitch excursion size and prominence”. *Journal of Phonetics*, Vol 13(3), 299-308.
- Risso, M. S., Silva, G. M. O., Urbano, H. (2006) “Traços definidores dos marcadores discursivos”. In Jubran, C. C. A. S. & Koch, I. G. V. (Org.) (2006). *Gramática do português culto falado no Brasil*. Volume 1: Construção do texto falado. Campinas: Editora da Unicamp.
- Rosenberg, A., Hirschberg, J. (2009) Detecting Pitch Accents at the Word, Syllable, and Vowel Level, NAACL/HLT, Boulder, CO.
- Saussure, F. et al. (1916) *Curso de linguística geral*. 24a. ed (2002). São Paulo, SP: Cultrix.
- Sagisaka, Y., Campbell, N., Higuchi, N. (1996) *COMPUTING prosody: computational models for processing spontaneous speech*. Coautoria de. New York, N.Y.: Springer, 1996.
- Saltzman, E. L. (1995). Dynamics and coordinate systems in skilled sensorimotor activity. In Port, R. & van Gelder, T., (Eds.). *Mind as motion: Dynamics, behavior, and cognition*. Cambridge, MA: MIT Press.

- Scherer, K. R. (1984) On the nature and function of emotion: a component process approach. In Scherer, K. R. Ekman, P. (Eds.) *Approaches to emotion*. Hillsdale, NJ: Lawrence Erlbaum. 1984. p. 293-318.
- Silva, M. C. P. S. & Crescitelli, M. F. C., (2006) “Interrupção”. In Jubran, C. C. A. S. & Koch, I. G. V. (Org.) (2006). *Gramática do português culto falado no Brasil*. Volume 1: Construção do texto falado. Campinas: Editora da Unicamp.
- Silva, P. (1998) *A Filosofia da Ciência de Paul Feyerabend*, Lisboa, Instituto Piaget.
- Silverman, K., M. Beckman, J. Pitrelli, M. Ostendorf, J. Pierrehumbert, J. Hirschberg, and P. Price (1992). TOBI: A Standard Scheme for Labeling Prosody. *Proceedings of the International Conference on Spoken Language*, Banff.
- Sonntag, G., P. & Portele, T. (1998) Comparative evaluation of synthetic prosody with the PURR method. *International Conference on Spoken Language Processing*.
- Tenani, L. E. (1996) Análise prosódica das inserções parentéticas no corpus do projeto da gramática do português falado. Dissertação de Mestrado. Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas.
- Thelen, E. and Smith, L.B. (1994) *A Dynamic Systems Approach to the Development of Cognition and Action*, MIT Press.
- van Gelder. & T. Port, R. (1995). It's About Time: An Overview of the Dynamical Approach to Cognition. In R. Port, & T. van Gelder, (Eds.). *Mind as motion: Dynamics, behavior, and cognition*. Cambridge, MA: MIT Press.
- Wightman, C.W. (2002): ToBI or not ToBI. In: *Proceedings of. Speech Prosody*, Aix-en-Provence, France, p. 25-29.
- Wong, S.W., Schreiner, C.E., 2003. Representation of CV-sounds in cat primary auditory cortex: intensity dependence. *Speech Communication*. 41, p. 93-106.
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics* 27, p.55-105.
- _____(2005). Speech melody as articulatorily implemented communicative functions. *Speech Communication* 46, p. 220-251.
- _____(2006). Speech prosody as articulated communicative functions. In *Proceedings of Speech Prosody 2006*, Dresden, Germany.
- _____(2010) In defense of lab speech. *Journal of Phonetics* 38: p. 329-336.

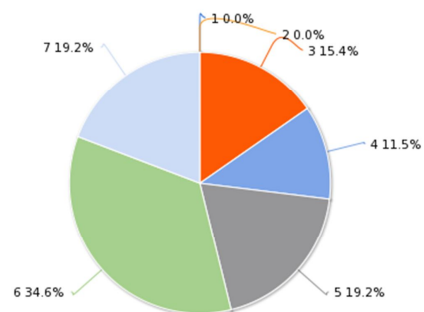
AnexoI

Summary Report - Apr 26, 2012

Survey: gênero/controle

1. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	0	0%
6	0	0%
5	4	15.4%
4	3	11.5%
5	5	19.2%
6	9	34.6%
7	5	19.2%

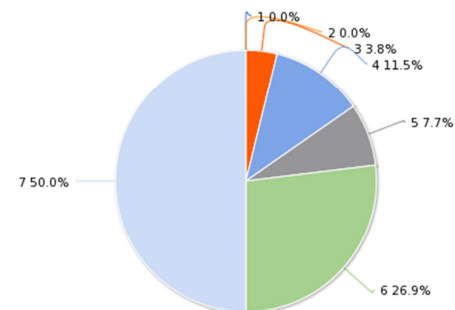
Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Entrevista oral pessoalmente

Count	Response
2	Explicação
1	Fala em exposição ou palestra.
1	Leitura
1	Resposta
1	conversa formal
2	Entrevista
1	entrevista de rádio
1	entrevista na televisão
13	Explicação
1	Explicação
1	Palestra

2. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	0	0%
6	0	0%
5	1	3.8%
4	3	11.5%
5	2	7.7%
6	7	26.9%

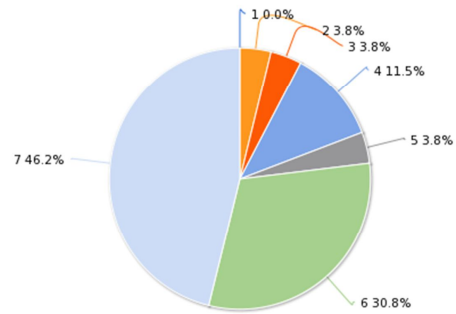
7	13	50%
---	----	-----

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
2	Conversa
1	Conversa informal
2	Entrevista
1	Explicação
1	Explicação
2	conversa
5	conversa informal
5	entrevista
1	entrevista de rádio
1	entrevista na televisao
2	explicação
2	narração
1	resposta à entrevista

3. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



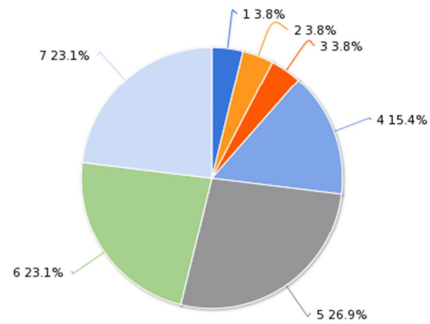
Value	Count	Percent %
7	0	0%
6	1	3.8%
5	1	3.8%
4	3	11.5%
5	1	3.8%
6	8	30.8%
7	12	46.2%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Conversa
1	Conversa Informal
1	Conversa informal
1	Diálogo informal
1	Entrevista
1	Explicação
1	argumentação
8	conversa informal
1	debate
1	debate na televisao
2	entrevista
4	explicação
1	exposição
2	palestra

4. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	1	3.8%
6	1	3.8%
5	1	3.8%
4	4	15.4%
5	7	26.9%
6	6	23.1%
7	6	23.1%

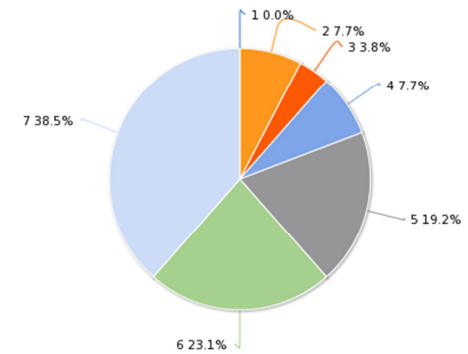
Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Conversa informal
2	Entrevista
2	Explicação
1	conversa informal
1	declaração em reportagem jornalística
6	entrevista
1	entrevista - explicação de fatos
1	entrevista de trabalho?
1	entrevistaa

Count	Response
7	explicação
2	narração
1	resposta

5. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	0	0%
6	2	7.7%
5	1	3.8%
4	2	7.7%
5	5	19.2%
6	6	23.1%
7	10	38.5%

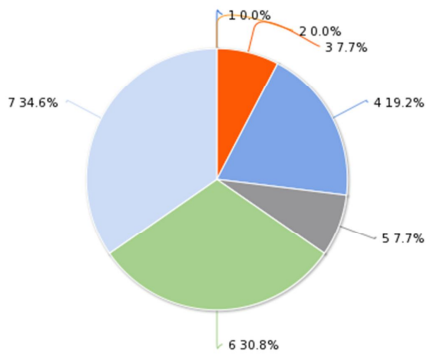
Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Conversa
4	Explicação
1	comentário

Count	Response
1	conversa formal
3	conversa informal
1	conversa informal - explicação
1	diálogo de rádio
12	explicação
1	leitura
1	programa de rádio

6. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



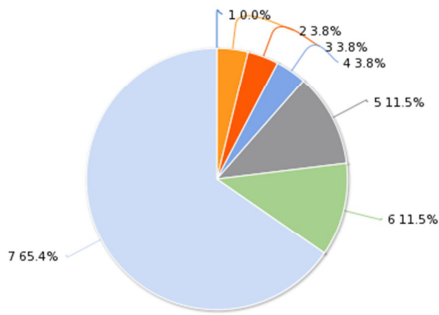
Value	Count	Percent %
7	0	0%
6	0	0%
5	2	7.7%
4	5	19.2%
5	2	7.7%
6	8	30.8%
7	9	34.6%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Comentário
1	Conversa
1	Conversa informal
2	Entrevista
1	Explicação
2	conversa
1	conversa com comentário
7	conversa informal
1	entrevista de tv/rádio
1	entrevista-explicação
1	explicação
1	fala da Gbriá Kalil rs
5	narração
1	reportagem de TV

7. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



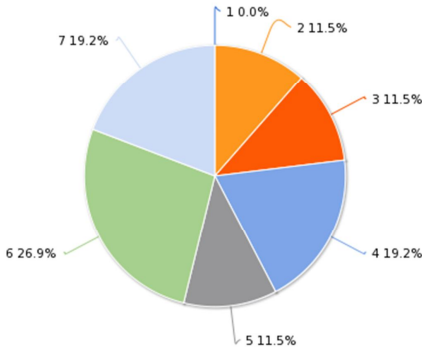
Value	Count	Percent %
7	0	0%
6	1	3.8%
5	1	3.8%
4	1	3.8%
5	3	11.5%
6	3	11.5%
7	17	65.4%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
2	Conversa informal
1	Entrevista
1	Explicação
1	Narração
1	Narração
1	c0nversa informal - narração
2	conversa
6	conversa informal
1	depoimento
1	entrevista
3	explicação
4	narração
1	narração - comédia?
1	reclamação

8. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	0	0%
6	3	11.5%
5	3	11.5%
4	5	19.2%
5	3	11.5%
6	7	26.9%
7	5	19.2%

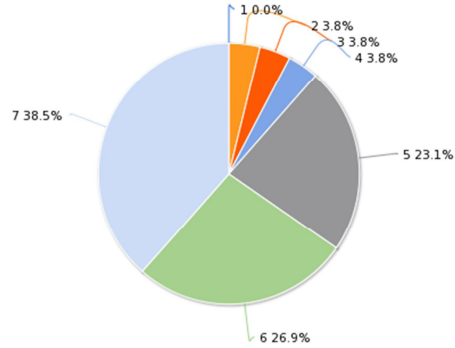
Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
3	Entrevista
1	Entrevista via telefone
1	Explicação
1	Narração
1	conversa
1	conversa por telefone
1	entrevista
7	entrevista

Count	Response
1	entrevista jogador de futebol para rádio
1	entrevista por telefone
1	entrevista-explicação
7	explicação

9. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	0	0%
6	1	3.8%
5	1	3.8%
4	1	3.8%
5	6	23.1%
6	7	26.9%
7	10	38.5%

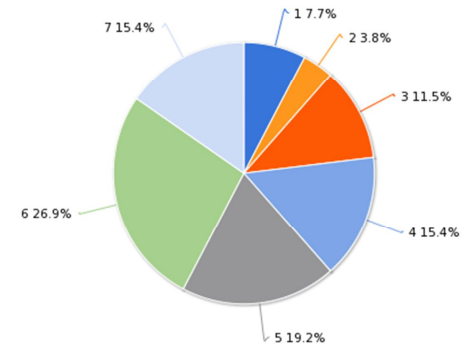
Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Conversa
2	Conversa informal

Count	Response
2	Entrevista
1	Entrevista oral pessoalmente
1	conversa
1	conversa formal
4	conversa informal
1	conversa/entrevista
2	entrevista
1	entrevista-explicação
3	explicação
5	narração
1	reportagem
1	resposta

10. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	2	7.7%
6	1	3.8%
5	3	11.5%
4	4	15.4%
5	5	19.2%

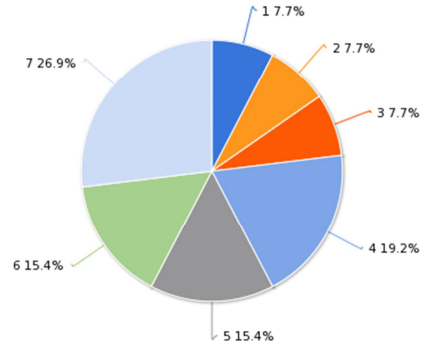
6	7	26.9%
7	4	15.4%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Conversa Informal
1	Explicação
2	Explicação
1	Leitura de "teleprompter"
3	conversa informal
4	entrevista
1	entrevista formal
7	explicação
1	leitura
1	locução de rádio
3	narração
1	recitação de dados decorados

11. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



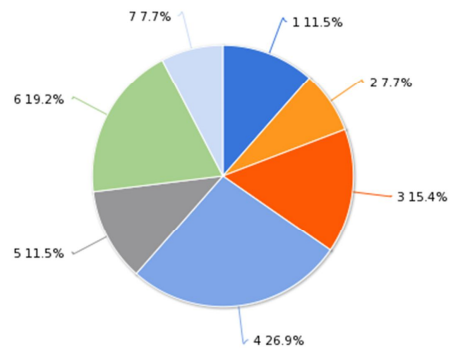
Value	Count	Percent %
7	2	7.7%
6	2	7.7%
5	2	7.7%
4	5	19.2%
5	4	15.4%
6	4	15.4%
7	7	26.9%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Conversa informal
1	Entrevista informal
1	Explicação
1	Explicação preparada
1	Narração
1	Opinião
3	conversa informal
2	entrevista
5	explicação
1	futuro aton!
1	história narrada
2	leitura
3	narração
1	narração - explicação
1	narração-leitura
1	resposta pesquisa de rua

12. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	3	11.5%
6	2	7.7%
5	4	15.4%
4	7	26.9%
5	3	11.5%
6	5	19.2%
7	2	7.7%

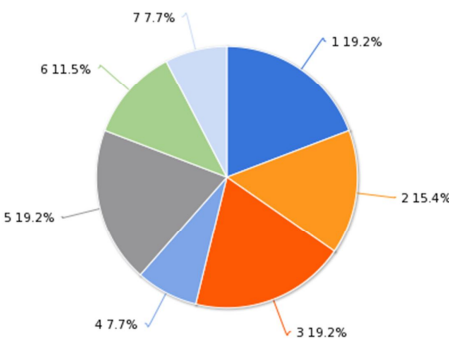
Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Explicação
1	Explicação
3	Narração
1	Palestra
1	aula
1	depoimento
2	entrevista
1	entrevista - explicação
9	explicação

Count	Response
1	leitura
4	narração
1	notícia

13. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	5	19.2%
6	4	15.4%
5	5	19.2%
4	2	7.7%
5	5	19.2%
6	3	11.5%
7	2	7.7%

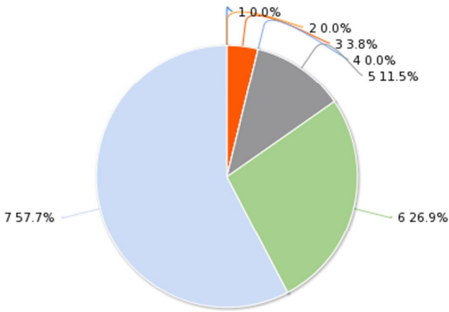
Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
2	Explicação
1	Explicação formal

Count	Response
2	Leitura
1	Palestra
1	entrevista
1	entrevista - explicação - leitura
10	explicação
1	explicação - aula
4	leitura
1	notícia lida
1	palestra
1	recitação de dados decorados

14. Quanto à espontaneidade, você considera esse exemplo de fala:

Copy of Quanto à espontaneidade, você considera esse exemplo de fala:



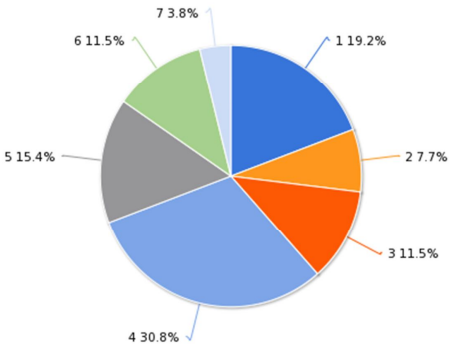
Value	Count	Percent %
7	0	0%
6	0	0%
5	1	3.8%
4	0	0%
5	3	11.5%
6	7	26.9%
7	15	57.7%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Conversa
1	Conversa Informal
2	Conversa informal
1	Entrevista
1	Narração
1	comentário
1	conversa
8	conversa informal
1	conversa informal - entrevista - narração
5	entrevista
4	narração

15. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	5	19.2%
6	2	7.7%

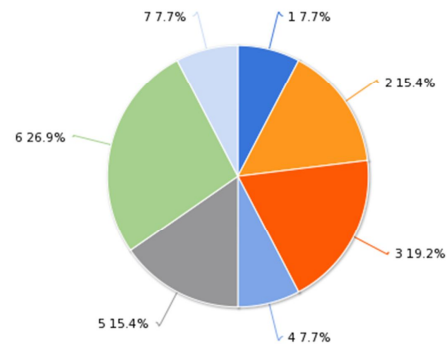
5	3	11.5%
4	8	30.8%
5	4	15.4%
6	3	11.5%
7	1	3.8%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
3	Explicação
2	Leitura
1	Palestra
1	entrevist - explicação
1	entrevista
9	explicação
1	explicação de aula
4	leitura
1	leitura-explicação
2	narração
1	palestra

16. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



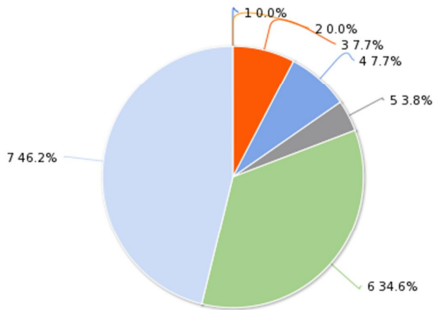
Value	Count	Percent %
7	2	7.7%
6	4	15.4%
5	5	19.2%
4	2	7.7%
5	4	15.4%
6	7	26.9%
7	2	7.7%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
4	Explicação
1	Leitura
1	Narração
1	conversa informal
3	entrevista
1	entrevista - explicação
9	explicação
1	leitura
2	narração
1	recitação de dados decorados
1	reportagem
1	resposta

17. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	0	0%
6	0	0%
5	2	7.7%
4	2	7.7%
5	1	3.8%
6	9	34.6%
7	12	46.2%

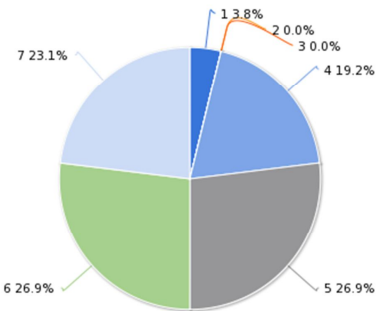
Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Comentário
1	Conversa
2	Conversa informal
1	Entrevista
1	Explicação
1	comentarista
1	comentários de um programa de esportes
3	conversa
9	conversa informal

Count	Response
1	conversa sobre tema - rádio ou tv
1	debate
2	entrevista
1	entrevista - explicação
1	explicação

18. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	1	3.8%
6	0	0%
5	0	0%
4	5	19.2%
5	7	26.9%
6	7	26.9%
7	6	23.1%

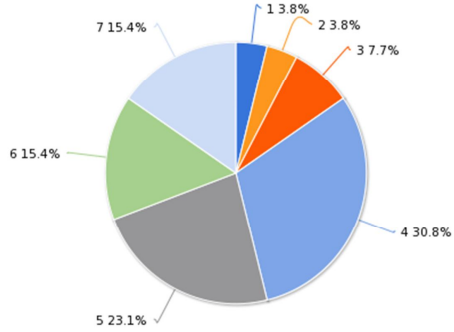
Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Comentário

Count	Response
1	Debate
1	Entrevista
1	Explicação
2	Explicação
1	comentário sobre uma notícia
1	conversa
2	conversa formal
2	conversa informal
1	conversa sobre tema - rádio ou tv
7	explicação
4	narração
1	opinião
1	reportagem

19. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	1	3.8%
6	1	3.8%
5	2	7.7%
4	8	30.8%

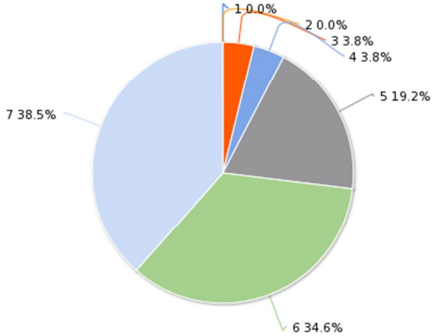
5	6	23.1%
6	4	15.4%
7	4	15.4%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Comentário
4	Explicação
1	Palestra
2	entrevista
14	explicação
1	narração
1	palestra
1	recitação de dados decorados
1	resposta

20. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	0	0%
6	0	0%

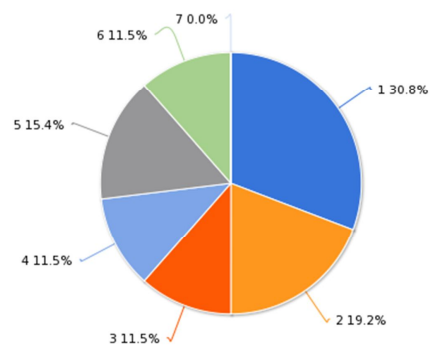
5	1	3.8%
4	1	3.8%
5	5	19.2%
6	9	34.6%
7	10	38.5%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
1	Conversa informal
3	Entrevista
2	Explicação
2	conversa informal
1	debate
5	entrevista
1	entrevista - explicação
1	entrevista de atleta
1	explicação
5	explicação
4	narração

21. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



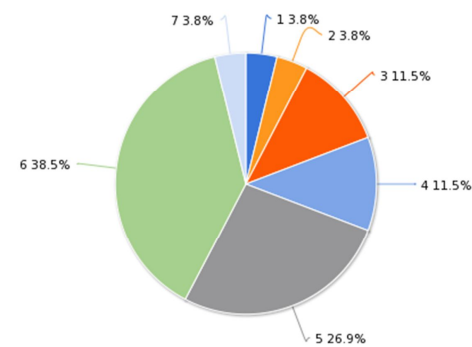
Value	Count	Percent %
7	8	30.8%
6	5	19.2%
5	3	11.5%
4	3	11.5%
5	4	15.4%
6	3	11.5%
7	0	0%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
5	Leitura
1	Leitura formal
7	explicação
1	explicação de aula
1	fala acadêmica
9	leitura
2	palestra

22. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
-------	-------	-----------

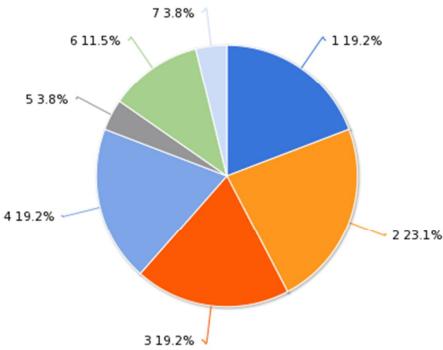
7	1	3.8%
6	1	3.8%
5	3	11.5%
4	3	11.5%
5	7	26.9%
6	10	38.5%
7	1	3.8%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
3	Entrevista
1	Entrevista formal
1	Explicação
1	Opinião
1	conversa formal
1	conversa informal
1	conversa sobre tema - rádio ou tv
4	entrevista
1	entrevista - opinião
1	entrevista-explicação
10	explicação
1	narração

23. Quanto à espontaneidade, você considera esse exemplo de fala:

Quanto à espontaneidade, você considera esse exemplo de fala:



Value	Count	Percent %
7	5	19.2%
6	6	23.1%
5	5	19.2%
4	5	19.2%
5	1	3.8%
6	3	11.5%
7	1	3.8%

Associe, livremente e de acordo com sua opinião, este exemplo a um tipo ou situação de fala (p.ex.: conversa informal, narração, leitura, explicação, etc...)

Count	Response
2	Explicação
3	Leitura
1	Palestra
1	aula expositiva
12	explicação
1	fala acadêmica
5	leitura
1	palestra

O teste terminou, obrigada! Por favor, preencha seus dados abaixo:

Nome completo: *omitidos*

Idade

Count	Response
1	20
2	21
1	22
3	23
2	25
3	26
1	27
2	28
2	29
3	30
1	31
1	33
1	34
1	39
1	43

E-mail: omitido

Anexo II

appleTV1

- 0 - 1 - (1)mas é:: o google TV NA VERDADE é uma quebra de paradigma um pouco diferente do Apple TV :::
(2)ah:: é baseada nnn em algo que o Apple TV já vinha fazendo
(3)que eh :: o consumo streaming de video
2 - (4)QUER DIZER você vai assisitr o video vindo sob demanda da internet no momento :: em que você está assistindo
(5)você num baixa mais :: armazena ele todo na sua casa :: e começa assistir...

appleTV2 (continuação do anterior, mas com outro falante)

- 0 - 1 - (1)pausa só pra todo mundo entender
2 - 3 - (2)você vai baixar um vídeo pra você assistir ::
(3)aí você baixa ele inteiro pela internet
(4)depois você bota um play lá : né [isso] pra tocar e passa
4 - (5)o streaming que ela tá falando é você vê ao mesmo tempo :::
(6)ce vê dois minutos ::: depois dele ser baixado
(7)entendeu?
(8)e conforme ce vai vendo isso ele vai sumindo

brasil_modaTR

- 0 - 1 - (1)eh eh uma coisa de ::: ehh :: de ser ::: eh : informal :: de ser leve de ser alegre de ser uma coisa
(2)e não é associado a luxo
2 - 3 - (3)veja bem nós temos uma imagem
(4)sobretudo na moda né? :::
4 - (5)eles admitem eles :: eles : OLHAM com prazer a nossa roupa esportiva :: biquini eh :
(6)tudo que é roupa informal
(7)nós não temos lugar ::: espaço
(8)não há expectativa [do terno e gravata]
5 - (9)AGORA o Brasil tem uma imagem tão positiva lá fora ::
(10)você pode fazer a experiência
(11)eu ja fiz essa experiência :
(12)todo mundo já fez :::
6 - (13)você chega lá fora oh ou você tá diante dum estrangeiro
(14)você diz que ce é brasileira ::
(15)todo mundo sorri :::
(16)imediatamente as pessoas sorriem [é verdade]
(17)vai dizer que ce é Belga ohh ohh qualquer coisa ninguém ri
7 - (18)ENTÃO : a imagem do Brasil : é uma imagem solar agradável simpática e tudo ::
(19)nós temos que aproveitar ess'imagem

- (20)e colocar esses elementos na nossa moda
- (21)pra qu'ela seja vista da mesma maneira

darwin

- 0 - 1 - (1)as outras visões ::: de criação do aparecimento das espécies ::: não são teorias científicas ::::
- (2)dessa forma é uma perda de tempo ficar comparando ::: o que seria uma teoria científica :::: com uma teoria :::: não científica
- 2 - (3)eu conheço ::: vários pesquisadores importantes ::: que ::: tem as suas crenças religiosas :::
- (4)contudo eles aceitam completamente a teoria evolutiva do Darwin :::
- (5)porque não :: consideram ::: qu'exista incompatibilidade entre as duas :::
- 3 - (6)as teorias são BASICAMENTE uma forma de ver o mundo ::
- (7)e as teorias científicas são :: interessantes do pon'de vista de que ::: elas tem em período de duração ::
- 4 - (8)nesse aspecto ::: a teoria do Darwin :: já tem cent'e cinquenta anos :::
- (9)isso é um verdadeiro record ::
- (10)ele praticamente só perde para teoria da gravidade de Newton

desfileTR

- 0 - 1 - (1)eh eh nn a brincadeira
- (2)QUER DIZER que a gente fala é que :: ehh o auê é muito maior do que ah :: ah : ah do que o faturamento vamos dizer ASSIM né?
- :::
- (3)o auê com a moda : é imenso ::: eh :: vende muito programa de rádio de televisão vende pro/ vende revista ::: eh
- 2 - (4) é um produto de entertainment da pesada moda
- (5)porque o povo :: ama : ver : desfile : é bonito mesmo desfile de moda : é bonito de ver :: é engraçado é interessante tudo:::
- 3 - (6)AGORA vende muito mais :: portanto :: entertainment do que vende roupa :
- (7)porque : não corresponde :: ah : a centimetragem né?
- (8)o o espaço que a moda ocupa ::: não corresponde à venda
- (9)não vende tanto quanto parece

desfile_melhorTR

- 0 - 1 - (1)OLHA : é o seguinte quando/ eu me lembro perfeitamente
- (2)aliás esse ano fo foi :: ah comemorou-se quinze anos de São Paulo Fashion Week ::
- (3)portanto né? faz faz ::: eh muito tempo [risos] ehhhh :: ehhhh melhorou muito ::
- 2 - (4)as empresas que se apresentavam primeiro copiavam muita roupa fora ::
- (5)a gente via no confronto na passarela :: uma duas três vezes a mesma roupa copiada fora
- 3 - (6)a partir do momento que ::: começou esse confronto e essa competição e essa visibilidade ::: houve um investimento enorme : em criatividade :: né? ::

- (7)houve um inv/ um investimento muito grande :: num desenvolvimento de :: produto
- (8) uma clara de entre/ tentativa de fazer uma moda brasileira ::
- 4 - (9)houve uma organização imensa do calendário de :: dos lançamentos :
- (10)organizou-se a indústria ::
- 5 - (11)porque antes todo mundo lançava cada um um pouco numa época :: ehh
- (12)não era uma coisa organizada não tinha um setor organizado ::
- 8 - (13)e a moda brasileira : de fato :: se beneficiou muitíssimo disso
- 6 - (14)é uma diferença brutal :: de qualidade de criatividade [uhum] de oferta ::
- 7 - (15)ah foi foi uma mudança grande

disco_livro

- 0 - 1 - (1)mais hoje em dia por exemplo eu to lançando o u::: um meu disco novo no formato dum livro também
- (2)ENTÃO vem o livro e o CD juntos
- (3)que é o Sambazz :::
- 2 - (4)e o Samba/ e o livro explica a produção do disco
- (5)ENTÃO além de ter o disco ainda tem o livro :
- 3 - (6)que eu escrevi todo ::: falando da produção do disco desd' a:: das composições até a finl/ finalização dele :::
- (7)também pra despertar um interesse na pessoa
- (8)ASSIM não simplesmente em baixar as o disco na in/ na internet de graça e :: ouvir as músicas ::
- (9)mas também pro cara : ter o:: ter a vontade de ler o livro ::: e:: e saber da história do disco né?

giseleTR

- 0 - 1 (1)mas não tem : uma Gisele Bundchen tem [] [] a Gisele Bundchen
- (2)num tem duas nem três tem a Gisele!
- 2 - (3)ENTÃO quando você chama uma : top desse jeito :: de fato aquilo levanta bastante o de/ o desfile
- (4)e faz com que :: a imprensa inteira esteja lá etc... ::
- (5)AGORA quando você chama um Lobão :: você distrai o OLHAr
- 3 - (6)eu fui ver por exemplo ah essa essa que ce tá falando essa essa coleção
- 4 - (7)que era da Reserva uma masc/ marca muculina :: Carioca ::
- 5 - (8)e eu também fiquei OLHAndo pro show do Lobão
- (9)eu nao prestei a menor atenção na roupa
- (10)nem lembro

industriaTR

- 0 - 1 - (1)as pessoas acham que com esse com esse aumento de consumo que teve especialmente da classe C D ::
- (2)que a indústria brasileira se beneficiou ::

- 2 - (3)e vou dizer que infelizmente não foi isso que aconteceu ::
3 - (4)o comércio sim :: e o consumidor sim tiveram mais oh mais : opções [aham]
(5)mas a indústria brasileira não,
4 - (6)por quê? porque nós não temos preço
5 - (7)ENTÃO essa roupa toda que veio pra abastecer :: com preço bom ::
(8)pra pra abastecer classe C e D :: veio da China e da Índia :::
6 - (9)ou seja não foi a indústria brasileira que se beneficiou disso

irradiado_contaminado

- 0 - 1 - (1)A: ser irradiado é diferente de ser contaminado?
(2)B: sim :::
(3)A: ser irradiado é como você falou : ser irradiado pela luz
(4)B:isso
2 - (5)A: e não necessariamente você se torna luminoso depois
(6)já contaminado você tem a presença do elemento radioativo/
(7)B: no seu corpo : dentro de você, na sua pele
3 - (8)se tá na sua pele ele vai acabar difundindo né? :::
(9)ENTÃO aí você passa a transportar né? aquele material com você :::
4 - (10)as vítimas do acidente de do no de Goiânia algumas se alimentaram preparando um sanduíche com mãos que tavam contaminadas com o pó né? de Césio cent'rinta e sete :::
5 - (11)e essas pessoas passaram a ser fontes ambulantes elas próprias irradiavam tudo ao redor ::
(12)mas quem afetava efetivamente a saúde delas era o componente Beta do Césio ::
(13)o Césio cent'rint'e sete emite :: Beta: e Gama ::
6 - (14)ENTÃO eh hh você estimar as doses é é possível tem toda instrumentação pra isso
(15)mas efetivamente eh eh eh algo complexo

jogador_futebol

- 0 - 1- (1)sem dúvida eu:: :::
(2)eu particularmente desde dois mionze :::
(3)eh foi muito :::
(4)eh muito bom no sentido :::::
(5)eh seleção né até porque:: ::
2- (6)ehh eu tive oportunidade de tá na ::: segundo goleiro terceiro goleiro
(7)mas :::: eh esse ano foi mei' stranho né
(8)graças a deus pude me afirmar né
(9)então ::: eh eh
3 - (10)mais seria :: com certeza bem melhor realmente uh : uh botafogo né
(11)realmente a gente conseguisse nossos objetivos
(12)que é o título né

4 - (13)mas particularmente foi muito bom esse dois mionze pra mim

jornalismo1

- 0 - 1 - (1)ah:::: eu acho que muitas matérias elas falam por si só né? ::
(2)que a gente não precisa voltar e fazer algum comentário em relação àquilo porque é uma indignação que é evidente :::
2 - (3)a gente quando colocou no ar esses três exemplos que o Galo por exemplo separou ::: ehh a gente colocou isso também lógico nos telejornais da Globo
(4)não precisa dizer mais nada ::
(5)acho que aquilo ali se eh uh as pessoas falam por si só :::
3 - (6)eu acho que a gente tem que ser imparcial :::
4 - (7)eh:: oh:: jornalismo tem isso né? a gente tem que dar : todos os lados né?
(8)a gente aprende na faculdade que :: tem dois lados da notícia
(9)eu digo que tem muito mais as vezes tem mm muito mais lados do que apen : apenas dois ::
(10)e a gente tem que mostrar todos esses lados da notícia :: pro nosso telespectador:::
(11)quem tem que fazer o julgamento são eles :: não sou eu

jornalismo2

- 0 - 1 - (1)acho que sim ::
(2)a gente consegue perceber isso pelos e-mails que a gente recebe :: ehh pelas ligações pro CAT
2 - (3)que é Central de Atendimento ao Telespectador :: da TV Globo
(4)eu acho que sim :::
(5)acho que acho que cada vez mais :: o nosso público :::
3 - (6)até porque a gente hoje em dia tem o acesso ah você tem internet ali na ss/ na na sua frente :::
(7)vinte quatro horas por dia sabe tudo que t'acontecendo no mundo inteiro :: você tem o mundo na sua frente no computador né? :::
4 - (8)isso (é o q/si) quanto mais informação :: acho que a pessoa consegue :: cada vez mais :: se preparar e e e e e e saber coisas:: pra poder julgar :: aquilo que ela tá vendo :: na televisão dela né? ::
5 - (9)ENTÃO ASSIM eu sinceramente acho que não cabe ::: a nós : âncoras
(10)acho que cabe ao público julgar aquilo
6 - (11)até porque :: eu tenho uma :: formação diferente eu tenho uma cultura diferente cada um tem né? :: a a o os seus princípios :::
(12)ENTÃO de repente uma coisa que me choca barbaramente a você pode não chocar :: vice-versa né? :: [uhum]

jornalismo3

- 0 - 1 - (1)eu acho que eu

- (2)eu acho que sim
(3)quando quando tem uma matéria eh muito forte de uma coisa muito absurda : eh num eh a minha cara diz tudo :::
2 - (4)acho q'eu sou uma pessoa bem expressiva ::
(5)ENTÃO eh eh não em palavras a gente não né? ::
3 - (6)tem essa coisa mesmo da imparcialidade de se preocupar com isso q'eu eu acho qu'é importante ::
4 - (7)mas eu acho que tem vezes que ASSIM que oh oh oh que a gente :: o o nosso rosto não esconde né? ::
(8)não tem jeito :::

jornalismo4

- 0 - 1 - (1)eh' NA VERDADE ASSIM não é um cara que dá opinião né?
(2)a gente ASSIM pra opinião a gente tem os nossos comentaristas nossos colunistas
(3)tem :: o Arnaldo Jabor, tem o Sardenberg falando de economia, tem o Eraldo Pereira falando de política :::
2 - (4)eh ah o o que mudou é que hoje em dia os âncoras eles são jornalistas mesmo eles não só apresentam o jornal eles editam o jornal
nós somos editores também :
3 - (5)ENTÃO ASSIM a gente tá no no :: oh trabalho braçal alí por trás também né?
(6)não é só :: a vitrine daquilo tudo tem um trabalho enorme de uma equipe enorme atrás da gente ::
(7)e a gente faz parte dessa equipe [uhum] né?
(8)o que a gente tá mostrando pro público a gente também :: tá tá fazendo parte dessa equipe ::
(9)acho que mais por aí ::
4 - (10)claro que existem alguns âncoras :: que :: são mais :: eh dão mais opinião né? ::
(11)mas NA VERDADE o papel do âncora não é esse
(12)esse é o papel de um colunista de um comentarista

lixo

- 0 - 1 - (1)eu vou falar um pouco do lixo : depois eu falo do esgoto ::::
2 - (2)no lixo gente pode fazer a geração d' energia :: a partir :: da recuperação do gás ::: ou da produção de combustíveis :::
3 - (3)esse gás ele é obtido :: quando da decomposição : da matéria orgânica dos restos de comida do papel presente no lixo :::
(4)isso pode acontecer nos aterros sanitários ::: ou : em : biodigestores :::
4 - (5) a grande diferença é a quantidade produzida e o : tempo :: de produção :::
5 - (6)os aterros produzem mais ::
(7)mas demoram ::: até oitenta anos
6 - (8)enquanto os biodigestores produzem menos
(9)mas em cerca de ehh duas semanas já te dão a quantidade disponível ::: pra poder tocar um projeto de energia
(10)depois a matéria orgânica que sobra :: digerida nos digestores el' é encaminhada pr' adubo : ou pra outras :: alternativas
7 - (11)enquanto no aterro : há necessidade de ficar um monitoramento :: por décadas
8 - (12)esse gás pode se utilizado ::: pra geração elétrica ::
(13)mas ele também pode ser usado pra fins industriais pra fins veiculares ::
(14)depende do nível de tratamento que se dê :: ao gás

loira_mtv

0 - 1 - (1)OLHA eu queria chorar porque eu não tenho ASSIM ahh condição ::: de prever o meu futuro :: além do ponto final se tornando uma outra língua brasileira :::

2 - (2)porque eu acho que o futuro ::: a Deus pertence :::: e Aton ::::

(3)ENTÃO : o futuro pertence ao ponto Aton ::::

3 - (4)porque eu acho que o nosso futuro é muito reluzente ::::

(5)que no futuro a gente pira ::::

(6)não vai ter mais a que ::: se reportar ::

4 - (7)ENTÃO as réplicas elas são constantes ::

(8)ce não pode ah ah adequar que a sua réplica um dia foi um ::: ma inocência né? ::::

(9)ENTÃO ::: os treplicas são aquelas crianças que vão defender o futuro Aton ::::

5 - (10)ai eu vou explicar o porque Aton :::::

6 - (11)porque ce nunca sabe de onde vem a morte ::

(12)as vezes você pode tá numa cozinha junto c'o cara :::: e achar que aquele lugar não existe mais porque você não pode voltar

(13)mas o lugar existe ::: ninguém vai derrubá-lo ::

7 - (14)ENTÃO ce nunca sabe da onde vem a morte e o futuro :::: reside :: em relutar :::::

(15)em fazer uma ::::: revolução interna :::: e conduzir até as urnas :: todo esse sacolejo dos anos que vem

medo

0 - 1 - (1)e essa superstição leva as pessoas a terem medo :: né?

2 - (2)esse medo como eu te disse vira mania

(3)essa mania vira TOC ::

3 - (4)ENTÃO por isso que muitas pessoas tem muitos me/ muito medo das coisas e as vezes não sabe onde vem :::

(5)e as vezes vem dessas pequenas coisas ::

4 - (6)eu não sei você mas eu tive um tempo na minha vida quando era mais nova a gente tinha muito contato com dono de fazenda fazendeiro essas coisa(s) ::

(7)e eles contavam aquelas coisas né dos escravos que arrastava corrente na madrugada

5 - (8)ENTÃO você ia p'essas fazendas pra dormi quem disse que ce dormia? :::

(9)se passava a noite inteira com medo que se ia ouvi corrente

(10)ai você acabava ouvindo

(11)porque o seu medo faz você ouvir :::: né o medo faz a gent/ :

6 - (12)se tem medo de escuro se você ficar cinco minuto(s) no escuro você ouve voz cê ouve telefone tocar você ouve te chama(r) cê ouve cê sente toca(r) em você ::

(13)tudo isso é gerado pelo seu medo ::

(14)não tem nada daquilo ali ::: né

7 - (15)ENTÃO quando eu comentei :: que nem o dia/ o demônio aguenta ::

(16)QUER DIZER :: nem o demônio aguenta deus : né

(17)ENTÃO deus ele é ah ah ah perfeição
(18)ENTÃO se a gente confia nele :: acho que cê não tem que ter medo das pessoas

reator_nuclear

0 - 1 - (1)ENTÃO NA VERDADE um reator nuclear é uma termoelétrica digamos metida a besta né?
(2)um pouco mais complexa mais complicada e que tem riscos específicos
(3)mas BASICAMENTE a idéia é a mesma :::
2 - (4)você tem uma caldeira com água e você precisa de uma fonte de calor ::
3 - (5)no caso do reator é a fissão do átomo que produz esse calor :::
(6)ah eh no caso de uma termoelétrica a gás a carvão a óleo a lenha é qualquer combustível né? :::
4 - (7)ENTÃO cê : aquece a água isso gera vapor que sopra uma turbina : gira né? um dínamo grande ::
(8)igual hum a gente tem num carro e com isso produz eletricidade :::
5 - (9)ENTÃO tem tem riscos específicos
(10)porque se você tem um' explosão de caldeira numa termoelétrica comum :: bom explodiu aliviou a pressão foi embora' o calor acabou né? :::
6 - (11)no caso do reator nuclear não né?
(12)o que a gente tá vendo lá no Japão :: ele não tá gerando mais nada né? de de energia
(13)mas como a a fissão continua gerando calor :: que precisa ser removido :
7 - (14)e foi a falha :: a incapacidade deles de fazer isso remover o calor :
(15)que fez com que :: o núcleo do reator começasse ah ah fundir né? parcialmente e liberar : radiação pro ambiente

superstição

0 - 1 - (1)porque : tem uma con/ existia uma convicção de que o mundo ia acabar :: né?
(2)no dia que concid/ coincidissem isso :::
2 - (3)só que ASSIM eh :: a superstição o que que a superstição é?
(4)ela nada mais é do que aquela velh'insegurança né?
(5)ENTÃO eu acredito em tudo :: eu todo mundo o que todo mundo me fala eu eu : coloco em :::
3 - (6)porque a a superstição ela vem :: de uma pessoa que num pensa :: né?
(7)se voce for parae e pensar como é que um número :: pode ter poder e força em cima de você? :::
4 - (8)AGORA o que qu' ele tem poder e força?
(9)porque na bíblia fica/ na bíblia sagrada católica apostólica romana fala :: que o número meia meia meia : é o número da besta
5 - (10)só que o número da best' ele num tem zero meia zero meia de dois mil e seis :
(11)é o número meia meia meia ::
(12)é diferente :: né? :::
6 - (13)ENTÃO é por causa disso que acredito que há :: essas superstições

teatrol

0 - 1 - (1)eu comecei a estudar teatro com dezessete anos :::
(2)e na casa das sete mulheres eu já tava com vinte sete :::
2 - (3)e nesse período : eh eu estudei continuo estudando até hoje
(4)mas eu fiz parte de algumas companhias teatrais
(5)mas eu só trabalhava com teatro eu nunca tinha feito nada na televisão nem nada no cinema ::
3 - (6)a casa das sete mulheres foi meu primeiro papel na televisão
(7)até ENTÃO eu só fazia teatro

tramaTR

0 - 1 - (1)CARA NA VERDADE ASSIM a a Trama ela ela ::: deu muito certo porque além d/ de ser uma gravador' independente ::: eh co' a mentalidade gravador' independente
(2)eh eh el' apostou em pessoas que o merca/ o grande mercado não tava querend' apostar
2 - (3)porque eu lembro que ah ah na época que eu retornei ao Brasil ::: o que realmente interessava pras grandes gravadoras ou era o sertanejo ::: ou o axé : ou o pagode :: qualquer uma :
(4)[qualquer onda né?] é
(5)qualq/ qualquer outra coisa fora desses três era muito difícil a gravadora apostar :::
3 - (6)ENTÃO o mercado independente começou a servir como u-uma uma válvula de escape pra quem quaria ouvir outras coisas e pra quem também tava fazendo outras coisas
4 - (7)ENTÃO ::: eh a Trama junto com outras gravadoras independentes começaram apostar em outras coisas
(8)não era só nessa coisa do filho de artista
5 - (9)porque ficou muito marcado na Trama isso :::
6 - (10)eu saí da Trama em dois mil e cinco
(11)mas alí no começo ficou muito marcado porque tinha eu tinha o Simoninha o João Marcelo
7 - (12)era e e continua sendo um dos donos ::: um dos donos da Trama e tal :::
8 - (13)eh mas eu acho que a importância ::: da n...da nossa geração e da Trama e dessa mentalidade do do mercado independente musical :: foi justamente abrir um espaço pra coisas que não tavam rolando
9 - (14)aí depois que esse mercad'independente abriu esse espaço ::: a e as gravadoras co/ eh começaram apostar em outras coisas também
(15)ENTÃO isso foi muito legal [elas faliram todas]
10 - (16)um/ a situação não tá fácil :::
(17)eu por emexplo quando saí da trama em dois mil e cinco aí você começa perceber como esse mercado realmente eh eh eh difícil
11 - (18)por conta da pirataria
(19)por conta ::: d'um atraso também das dos próprios executivos das gravadoras
(20)que demoram muito pra en/ entender certos assuntos como a internet ::: eh como a venda online essas coisas

usp

0 - 1 - (1)eh eh eu f/ eu fico eu passo alí perto da USP quase todo dia né? :

- 2 - (2)eh ::: os caras alí aconteceu uma coisa interessante aí do estado brasileiro né? :
- (3)pra/ eh p/ por conta do do Serra também :: a gente tem um governador novo né?
- 3 - (4)ninguém sabe quem ele é :::
- (5)mas eh BASICAMENTE o que ele tem feito hoje em dia é não fazer nada :::
- 4 - (6)e a gent' também tem um reitor na USP que é mais ou menos o mesmo mesmo pacote ::
- 5 - (7)aí aconteceu uma coisa interessante essa semana que foi a privatização do portão da USP né? :::
- (8)porque né? a USP tem vários portões tem o portão principal :::
- 6 - (9)o estado como não tá muito interessado com a USP e o reitor parece também não tá nem mais na USP ::
- (10)eles pegaram tem um portão né?
- (11)tem os grevistas
- 7 - (12)tem uns oitenta grevistas que representam a class' inteira ::: e os caras ficam abrindo e fechando portão né? :::
- 8 - (13)e aí TIPO ASSIM cê vai entrar na USP cê tem que meio que negociar c'os caras né? :::
- (14)porque né? ::
- (15)foi privatizado né?
- 9 - (16)e o pessoal fala pô e ó :::: governo deixou explorar isso aqui né? :::
- 10 - (17)é mais ou menos como se dessem as ruas pro pessoal pros flanelinhas explorarem assim né? cê vai lá ::
- 11 - (18)eu fiquei pensando né? cara eh :: os caras podiam dar pelo menos um atestado pra eu levar pro empregador
- (19)ASSIM fala ó
- (20)eu vim pe/ eu vim pela USP daí o pessoal não me deixou passar :: eles me deram esse atestado falando que eu vou levar duas horas a mais ::
- (21)vou ter que dar volta inteira na USP pra chegar lá cara :::
- 12 - (22)que é que é uma bizarrice né?
- 13 - (22)tem oitenta pessoas que fecham o portão ninguém faz nada ninguém entra na USP faz ::
- (24)isso t' acontecendo faz três semanas
- (25)o estado não faz nada ::
- (26)o reitor não tá na USP::
- (27)e tá tudo certo

vergonha

- 0 - 1 -(1)ENTÃO eh duas coisas em relação aos ouvintes que mandaram qu/ é milhões de críticas em respeito ao ao Renan
- 2 - (2)uma coisa que eu discordo veement/ veementemente é eu não acho que eu tenho vergonha de ser brasileiro como muitos dos ouvintes começaram colocar
- 3 - (3)e outro qu'eu não se não me engano foi a um uma ouvinte de nome Raquel ::: falando que ela não sabe mais como cont/ : como falar pros filhos dela :: que vale a pena ser honesto no Brasil
- 4 - (4)pô eu acho que AGORA é o melhor momento de explicar prum filho que vale a pena ser honesto
- 5 - (5)a gent'em os exemplos claros de as pessoas que não são honestas no Brasil
- (6)é só mostra na televisão mostrar foto no jornal e ver o qu'elas estão fazendo ::
- 6 - (7)é o contraponto perfeito o exemplo perfeito ::
- (8)é o exemplo negativo as vezes eh melhor do qu'exemplo positivo
- 7 - (9)a gente tem que aproveitar isso pra :: pra mudar o país

(10)não pra continuar nessa nesse marasmo né?

violencia

- 0 - 1 - (1)eu acho que ainda existe um pouco de confusão [uhum] da população como toda lei ao ser :::: ah entrar em vigor ::::
(2)porque :: os casos da lei Maria da Penha são relacionados à violência doméstica :::::
2 - (3)ENTÃO pras pessoas entenderem a caracterização do que é uma violência doméstica ::: eles acabam confundindo
(4)porque a lei :: ela dá margem :::: algumas interpretações
(5)mas ao mesmo tempo ela está sendo mais a/ :::: aplicada pra lei ::: pra violência doméstica
3 - (6)que que seria?
(7)pessoas que coabitam :::: numa relação ou de casamento ou de : amasios enfim ::::: de parentesco # ahh e que sofre a violência::::
4 - (8)não adianta chegar c'uma violência de briga de duas vizinhas # que isso num vai se encaixar na lei [uhum]
5 - (9)de primeiro momento houve uma enxurrada né? de :::: de :::: procura :::
(10)mas a maioria dos casos não se a/ ::: não se aplicavam #
6 - (11)ENTÃO a a sociedade também as mulheres pra pras :: pra uqe elas possam sentir que a lei está realmente valendo existindo ::::
existe uma necessidade pra que elas voltem a procurar :::: de ser feit' uma triagem um direcionamento
7 - (12)qual que é o caso? :::
(13)porque as vezes quando ela chega na delegacia ela vai achando que a o caso dela se aplica :::
(14)só que muitas vezes não é caso de fazer um boletim de ocorrência :::
8 - (15)se NA VERDADE ela quer a separação que iss/ a
(16)a vítima tem que ter isso muito claro né ::::: se ela quer a separação ::
9 - (17)muitas vezes ela sofre a violência e ela não quer ::::
10 - (18)como que nós poderíamos ajudar essa vítima #
(19)a a encaminhando ela pr'um trabalho
(20)dando uma consciência de que há necessidade de ter um :: retorno econômico pra ela poder se auto ::: sustentar :::::
11 - (21)e com esse trabalho nós vamos poder também fazer um levantamento do número exato de mulheres ::::: ahhh que passarão por lá né?
:::: vítimas de violência [que são vítimas né?] é? ::
(22)e até pra poder estudar o perfil do agressor né? [tá certo]
12 - (23)que até ENTÃO era difícil

Anexo III


```

# Name:      : F0_extrema.praat
# Version    : 0.10 (beta)
# Author     : Pablo Arantes <parantes(at)fastmail.fm>
# Created    : 25.11.2008
# Last change : 21.03.2009
#
# Purpose:
# -----
# Finds extreme points (local or global maxima and minima) in
# smoothed F0 and F0 velocity (first discrete derivative) contours.
#
# Input:
# -----
# A Sound or Pitch object must be loaded to the Objects list
# and selected. Also a TextGrid having the same name as the
# Sound file must be readable from a folder specified by the
# user at the script's GUI.
#
# Output:
# -----
# TextGrid with F0 and F0 velocity extreme points marked in two
# separate point tiers and one interval tier with user-provided
# segmentation. Optionally it is possible to save intermediate
# results (raw and smoothed interpolated F0 contours, F0 velocity
# contour and extreme point marks for F0 and F0 velocity) in a
# spreadsheet-like text file.
#
# Requires:
# -----
# A Sound and a TextGrid with at least one interval tier containing
# a user-defined segmentation
#
# Comments:
# -----
# Requires at least version 5.0.43 of Praat
#
# Copyright (C) 2008 Pablo Arantes
#
# This program is free software; you can redistribute it and/or modify
# it under the terms of the GNU General Public License as published by
# the Free Software Foundation; either version 2 of the License, or
# (at your option) any later version.
#
# This program is distributed in the hope that it will be useful,
# but WITHOUT ANY WARRANTY; without even the implied warranty of
# MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
# GNU General Public License for more details.
#
# You should have received a copy of the GNU General Public License
# along with this program. If not, see <http://www.gnu.org/licenses/>.

form Find extreme points in a F0 contour
sentence Grid D:\corpora_prosodia\celio\patarata\textgrids\
natural Segmentation 1
comment F0 smoothing factor
real Factor_ (Hz) 4
comment Smooth F0 derivative contour?
boolean Smooth_derivative 0
real Factor2_ (Hz) 3.5
comment Threshold method to be used to extreme point appending
optionmenu Threshold: 1
    option Glissando Threshold
    option 1 semitone
    option None
comment In case of Glissando threshold, choose a value
real Gtr 0.16
comment Collapse S chains?
boolean Collapse_chains 0
comment Look for outlying values in the F0 contour?
boolean Purge 0
comment Save script report?
boolean Script_report 1
comment Draw smoothed F0 and F0 1st derivative contours?
boolean Draw 1
comment Save TableOfReal as spreadsheet file?
boolean Save_table 0
comment Save F0 velocity contour as PitchTier?
boolean Save_PitchTier 0

```

```

        sentence Save_dir D:\work\
        comment Merge segmentation and F0 extremes tiers?
        boolean Merge_tiers 1
endform

stopwatch

! Terminate script execution if there are no selected objects
! or if more than one object is selected
n = numberOfSelected ()
if n <> 1
    exit One object must be selected, either a Pitch or a Sound.
endif

! Check what the selected object is
file$ = selected$ ()
type$ = extractWord$ (file$, "")

! Se for Sound, extrai-se Pitch a partir dele
if type$ = "Sound"
    sound = selected ("Sound")
    name$ = selected$ ("Sound")
    original_pitch = To Pitch... 0 75 600
else
    original_pitch = selected ("Pitch")
    name$ = selected$ ("Pitch")
endif

! Important info that will be used later on
file_end_time = Get end time

if purge = 1
    call purge_f0 original_pitch sound
else
    clearinfo
endif

call interpolate original_pitch

if smooth_derivative = 0
    call differentiate table 3 4 F0-vel 0 factor2
    subtitle$ = "%F_0 velocity"
else
    call differentiate table 3 4 F0-vel 1 factor2
    subtitle$ = "Smoothed %F_0 velocity"
endif

! Find F0 maxima
call extrema table 3 5 F0-extr
! Find F0' maxima
call extrema table 4 6 F0-vel-extr

! Index table by column 1 (time)
! It will be necessary later
call index_table table 1

!! Apply thresholds to raw extreme points

if threshold <> 3 ; only apply if threshold method is not "None"
    call filter_extreme table 3 5 1 6 threshold
    if collapse_chains = 1
        call collapse_chains table 5
    endif
endif

call control table 5 6

! Cria TextGrid que receberá marcação dos pontos extremos
extrema_grid = Create TextGrid... 0 file_end_time "F0 F0-velocity" F0 F0-velocity
! Renomeia TextGrid para 'name$'-extr
Rename... 'name$'-extr

if control.h > 0
    select table
    f0_h = Extract rows where column... 5 "equal to" 1
    call append_points f0_h 1 3 extrema_grid 1 H
endif
if control.l > 0

```



```

        select table
        f0_l = Extract rows where column... 5 "equal to" -1
        call append_points f0_l 1 3 extrema_grid 1 L
    endif
    if control.s > 0
        select table
        f0_s = Extract rows where column... 5 "equal to" 2
        call append_points f0_s 1 3 extrema_grid 1 S
    endif
    if control.r > 0
        select table
        f0_vel_r = Extract rows where column... 6 "equal to" 1
        call append_points f0_vel_r 1 4 extrema_grid 2 R
    endif
    if control.f > 0
        select table
        f0_vel_f = Extract rows where column... 6 "equal to" -1
        call append_points f0_vel_f 1 4 extrema_grid 2 F
    endif

    if draw = 1
        call draw_results
    endif

    !! Cleaning up the Objects list
    if control.h > 0
        select f0_h
        Remove
    endif
    if control.l > 0
        select f0_l
        Remove
    endif
    if control.s > 0
        select f0_s
        Remove
    endif
    if control.r > 0
        select f0_vel_r
        Remove
    endif
    if control.f > 0
        select f0_vel_f
        Remove
    endif

    if save_PitchTier = 1
        ! Cria objeto PitchTier que receberá o contorno
        f0_der_pt = Create PitchTier... 'name$'_df0-dx 0 file_end_time
        select table
        nrows = Get number of rows
        ! Atribui os pontos da TableOfReal ao objeto PitchTier
        for row to nrows
            select table
            time = Get value... row 1
            f0 = Get value... row 4
            select f0_der_pt
            Add point... 'time' 'f0'
        endfor
        Write to PitchTier spreadsheet file... 'save_dir$' 'name$'_df0-dx.PitchTier
        Remove
    endif

    if save_table = 1
        select table
        table_name$ = selected$ ("TableOfReal")
        Write to headerless spreadsheet file... 'save_dir$' 'table_name$'_table.txt
    endif

    if merge_tiers = 1
        original = Read from file... 'grid$' 'name$'.TextGrid
        interval_tier = Extract tier... 'segmentation'
        segm_grid = Into TextGrid
        select segm_grid
        plus extrema_grid
        Merge
        merged_grid = Rename... 'name$'-extr
        ! Cleaning up
    endif

```

```

        select original
        plus interval_tier
        plus segm_grid
        plus extrema_grid
        Remove
        extrema_grid = merged_grid
endif

clock = stopwatch

call write_report

select table
Remove

select extrema_grid

! -+--+ > > > Procedures < < < -+--+

procedure purge_f0 .f0_id .audio
# == Arguments ==
# .f0_id :: integer
# Numerical ID of the Pitch object being purged
#
# .audio :: integer
# numerical ID of the Sound object being analysed
#
# == Purpose ==
# The purpose of the procedure is to remove outlying pitch points in the
# F0 contour the script will work on. First it will extract the F0
# contour in a two-pass operation and then prompt the user to inspect
# the Pitch object and remove or add pitch points as s/he sees fit. When
# the user is done the script's execution will continue.
#
# The F0 extraction is a two-pass operation. The relevant parameters
# the procedure will manipulate are floor and ceiling F0 values. In the
# first pass the Pitch object is extracted using the default values (75
# and 600 Hz). In the second pass another Pitch object is extracted using
# optimal values for floor and ceiling estimated from the first Pitch
# object. The values are obtained using the following formulae:
#
# - F0 ceiling = 1.5*q3
# - F0 floor = 0.65*q1
#
# where q1 and q3 are respectively the first and third quartiles of the
# first Pitch object F0 values. This heuristic is suggested by Hirst
# [cf. D. Hirst, Proc. XVth ICPhS, Saarbrücken, 1233 (2007)]. Actually,
# Hirst suggests 0.75 as a coefficient for q1, but in my empirical
# experience 0.75 led to values slightly higher than the first-pass
# minimum value in situations where it was bona fide.

        select .f0_id
        .pitch_name$ = selected$("Pitch")
        .min = Get minimum... 0 0 Hertz None
        .min = round(.min)
        .max = Get maximum... 0 0 Hertz None
        .max = round(.max)
        .q1 = Get quantile... 0 0 0.25 Hertz
        .q1 = floor(.q1)
        .q3 = Get quantile... 0 0 0.75 Hertz
        .q3 = ceiling(.q3)
        .floor = floor(0.65*.q1)
        .ceiling = ceiling(1.5*.q3)
        select .f0_id
        Remove
        select .audio
        Edit
        editor Sound '.pitch_name$'
            Pitch settings... .floor .ceiling Hertz autocorrelation speckles
        endeditor
        select .audio
        ! Second pass F0 extraction
        .second_pass = To Pitch (ac)... 0 .floor 15 yes 0.03 0.45 0.01 0.35 0.14 .ceiling
        .voiced_before = Count voiced frames
        ! Prompt user to remove or add points in the 2nd pass Pitch object
        Edit
        editor Pitch '.pitch_name$'
            beginPause ("Unvoice spurious pitch points. (Selection > Unvoice)")

```

```

        .action = endPause ("Done", 1)
    if .action = 1
        Close
    endif
endeditor
.new_min = Get minimum... 0 0 Hertz None
.new_min = round(.new_min)
.new_max = Get maximum... 0 0 Hertz None
.new_max = round(.new_max)
.voiced_after = Count voiced frames
.balance = .voiced_after - .voiced_before
! Restoring object ID reference
original_pitch = .second_pass
editor Sound '.pitch_name$'
    Close
endeditor
! Writing information to the the info window
clearinfo
printline F0 extraction report'newline$'-----
printline 1st pass > minimum: '.min' Hz - maximum: '.max' Hz
printline estimated parameters: floor: '.floor' Hz - ceiling: '.ceiling' Hz
printline 2nd pass > minimum: '.new_min' Hz - maximum: '.new_max' Hz
printline net change: '.balance' points'newline$'
endproc

procedure interpolate .pitch_id

! Collecting information about raw F0 contour
! - time of first voiced frame
! - time of last voiced frame
! - duration of sound file
! The idea is to analyse only the voiced part, i.e.
! to trim the initial and final unvoiced parts
! of the sound file
!
select .pitch_id
.original_pitch_tier = Down to PitchTier
.n_of_points = Get number of points
.first_point = Get time from index... 1
.last_point = Get time from index... .n_of_points
.file_end_time = Get total duration

! Interpolate unvoiced parts of original PitchTier
! It will interpolate every point in the PitchTier,
! what is not good, because now the PitchTier is
! oversampled
Interpolate quadratically... 4 Hz
! Converting the PitchTier back to a Pitch object
! solves the problem: the new Pitch is sampled
! at a time frame of 10 ms again
.interpolated_pitch = To Pitch... 0.01 60 400
! Pitch is once more converted to a PitchTier for
! the trimming operation (couldn't find a way to do the
! trimming using the Pitch object).
.trimmed_pitch_tier = Down to PitchTier
! Now the undesired initial and final parts are trimmed
Remove points between... 0 .first_point
Remove points between... .last_point .file_end_time

! Trimmed PitchTier object is converted into
! a TableOfReal object, which will be the main
! working table for the script; the variable
! holding it's numerical ID is not prefixed
! with the dot for easier reference outside
! this procedure
!
table = Down to TableOfReal... Hertz

! Smooth the interpolated Pitch object using
! the user-specified bandwidth stored as 'factor'
! variable (3rd GUI form field)
!
select .interpolated_pitch
.smoothed_pitch = Smooth... factor

! Write the smoothed F0 values into the main TableOfReal
! Values go into column 3 which receives the label F0-sm
!

```

```

select table
.nrows = Get number of rows
Insert column (index)... 3
Set column label (index)... 3 F0-sm
for .row to .nrows
    select table
    .time = Get value... .row 1
    select .smoothed_pitch
    .f0_sm = Get value at time... .time Hertz Nearest
    select table
    Set value... .row 3 .f0_sm
endfor

! Cleaning the temporary objects not longer necessary
select .original_pitch_tier
plus .interpolated_pitch
plus .trimmed_pitch_tier
plus .smoothed_pitch
Remove

endproc

procedure differentiate .table .source_col .target_col .target_col_name$ .smooth .factor
select .table
.nrows = Get number of rows
Insert column (index)... .target_col
Set column label (index)... .target_col '.target_col_name$'
for .row from 2 to .nrows
    .prior = Get value... .row-1 .source_col
    .actual = Get value... .row .source_col
    .delta = .actual - .prior
    Set value... .row .target_col .delta
endfor
! Row 1, which is by definition 0, receives the value of row 2
.row2 = Get value... 2 .target_col
Set value... 1 .target_col .row2

! Handles smoothing of dF0/dx contour if the option is called for
if .smooth = 1
    .temp_ptier = Create PitchTier... sm_der_temp 0 file_end_time
    select .table
    call min table .target_col
    .height = 100 + abs(min.min)
    .nrows = Get number of rows
    ! Atribui os pontos da TableOfReal ao objeto PitchTier
    for .row to .nrows
        select .table
        .time = Get value... .row 1
        .point = Get value... .row .target_col
        .point = .point + .height
        select .temp_ptier
        Add point... '.time' '.point'
    endfor
    .temp_pitch = To Pitch... 0.01 40 400
    .temp_sm_pitch = Smooth... .factor
    for .row to .nrows
        select .table
        .time = Get value... .row 1
        select .temp_sm_pitch
        .point = Get value at time... .time Hertz Nearest
        .point = .point - .height
        select .table
        Set value... .row .target_col .point
    endfor
    select .temp_ptier
    plus .temp_pitch
    plus .temp_sm_pitch
    Remove
endif

endproc

procedure extrema .table .source_col .target_col .target_col_name$
select .table
.nrows = Get number of rows
Insert column (index)... .target_col
Set column label (index)... .target_col '.target_col_name$'
for .row from 2 to .nrows-1
    .previous = Get value... .row-1 .source_col
    .actual = Get value... .row .source_col

```

```

        .next = Get value... .row+1 .source_col
        if (.actual > .previous) and (.actual > .next)
            .extremum = 1
        elseif (.actual < .previous) and (.actual < .next)
            .extremum = -1
        else
            .extremum = 0
        endif
        Set value... .row .target_col .extremum
    endfor
endproc

procedure index_table .table .col
    select .table
    .nrows = Get number of rows
    for .row to .nrows
        Set row label (index)... .row '.row'
    endfor
endproc

procedure filter_extreme .table_id .value_col .ext_col .time_col .vel_ext_col .method
! = Procedure's parameters =
!
! .table_id :: integer
! numerical ID of the TableOfReal containing the script's main data
!
! .value_col :: integer
! number of the column containing the smoothed F0 values
!
! .ext_col :: integer
! number of the column containing boolean information on whether
! a F0 sample is an extreme
!
! .time_col :: integer
! number of the column containing the time of each F0 sample
!
! .vel_ext_col :: integer
! column containing boolean info on whether a F0-vel sample is an extreme
!
! .method :: integer
! Specifies what extreme filtering method was chosen by the user
!
    select .table_id
    .filter = Extract rows where column... .ext_col "not equal to" 0
    .nrows = Get number of rows
    for .row to .nrows
        if .row = 1
            select .filter
            .end_F0 = Get value... .row .value_col
            .end_time = Get value... .row .time_col
            .ext_type = Get value... .row .ext_col
            select .table_id
            .temp_table = Extract rows where column... .time_col "less than" .end_time
            if .ext_type = 1 ; in case it's a H extreme
                call min .temp_table .value_col
                .begin_F0 = min.min
                .begin_time = Get value... min.index .time_col
                .dur = .end_time - .begin_time
            else ; in case it's a L extreme
                call max .temp_table .value_col
                .begin_F0 = max.max
                .begin_time = Get value... max.index .time_col
                .dur = .end_time - .begin_time
            endif
            select .temp_table
            Remove
        else ; from row 2 up to the last
            select .filter
            .end_F0 = Get value... .row .value_col
            .end_time = Get value... .row .time_col
            .begin_F0 = Get value... .row-1 .value_col
            .begin_time = Get value... .row-1 .time_col
            .dur = .end_time - .begin_time
        endif
        ! Apply the chosen threshold methods
        if .method = 1 ; Glissando threshold method
            .delta_st = abs(hertzToSemitones(.end_F0) - hertzToSemitones(.begin_F0))
            .gtr = gtr
        endif
    endfor
endproc

```

```

        .threshold = .gtr / (.dur^2)
        .threshold = .threshold * .dur
        if .delta_st < .threshold
            .keep = 0
        else
            .keep = 1
        endif
    elseif .method = 2 ; +1 semitone threshold method
        .delta_F0 = abs(.end_F0 - .begin_F0)
        .threshold = semitonesToHertz ( hertzToSemitones(.begin_F0) + 1 ) - .begin_F0
        if .delta_F0 < .threshold
            .keep = 0
        else
            .keep = 1
        endif
    endif
    ! Change the extreme status if it doesn't pass the threshold criterion
    if .keep = 0
        ! Set .ext_col value to 2 - (S)tatic extreme
        select .filter
        .line$ = Get row label... .row
        select .table_id
        Set value... ' .line$' .ext_col 2
    endif
endfor

! = F0 velocity extrema removal =
! Velocity extremes occurring before F0 extremes
! that are marked as S extremes are removed

! .filter table has to be extracted again
! so that the changes are updated
select .filter
Remove
select .table_id
.filter = Extract rows where column... .ext_col "not equal to" 0
.nrows = Get number of rows
select .table_id
.filter_vel = Extract rows where column... .vel_ext_col "not equal to" 0
.nrows_vel = Get number of rows
for .row from 1 to .nrows
    select .filter
    .actual_ext = Get value... .row .ext_col
    if .actual_ext = 2 ; i.e. it's a S extreme
        .time_actual = Get value... .row .time_col
        if .row = 1
            select .table_id
            .time_prior = Get value... 1 .time_col
        else
            select .filter
            .time_prior = Get value... .row-1 .time_col
        endif
        for .row_vel to .nrows_vel
            select .filter_vel
            .vel_time = Get value... .row_vel .time_col
            ! All F0' extremes between the last and the
            ! present F0 extreme (marked as S) are unmarked as
            ! a F0' extreme
            if (.vel_time <= .time_actual) and (.vel_time > .time_prior)
                ! Set velocity extreme markings in the interval to 0
                select .filter_vel
                .line$ = Get row label... .row_vel
                select .table_id
                Set value... ' .line$' .vel_ext_col 0
            endif
        endfor
    endif
endfor

endfor
! Clean the clutter
select .filter
plus .filter_vel
Remove

endproc

procedure collapse_chains .main_table .ext_col

    select .main_table

```

```

.original = Extract rows where column... .ext_col "not equal to" 0
.nrows = Get number of rows
! Only proceed if there are at least 3 lines in the table
if .nrows >= 3
    for .row from 1 to (.nrows-2)
        select .original
        .ext = Get value... .row .ext_col
        if .ext = 2
            .ahead_1 = Get value... .row+1 .ext_col
            .ahead_2 = Get value... .row+2 .ext_col
            if (.ahead_1 = 2) and (.ahead_2 = 2)
                ! Set variables
                .begin_row = .row
                .end_row = .row + 2
                .seek = .row + 3
                .ahead_n = 2
                while (.seek <= .nrows) and (.ahead_n = 2)
                    .ahead_n = Get value... .seek .ext_col
                    if .ahead_n = 2
                        .end_row = .seek
                    endif
                    .seek += 1
                endwhile
                ! Collapse S extremes between .begin_row and .end_row not
                for .line from (.begin_row + 1) to (.end_row - 1)
                    select .original
                    .line$ = Get row label... .line
                    select .main_table
                    Set value... '.line$' .ext_col 0
                endfor
                ! The for loop continues from here
                .row = .seek - 1
            endif
        endif
    endfor
endif
select .original
Remove

endproc

procedure control .main_table .ext_col .vel_ext_col

! F0 extreme control
.h = 0
.l = 0
.s = 0
select .main_table
.control = Extract rows where column... .ext_col "not equal to" 0
.nrows = Get number of rows
for .row to .nrows
    .ext = Get value... .row .ext_col
    if .ext = 1
        .h += 1
    elseif .ext = -1
        .l += 1
    elseif .ext = 2
        .s += 1
    endif
endfor
select .control
Remove

!! F0' extreme control
! counts the number of "R" extremes (marked as 1 in .vel_ext_col)
.r = 0
! counts the number of "F" extremes (marked as -1 in .vel_ext_col)
.f = 0
select .main_table
.control = Extract rows where column... .vel_ext_col "not equal to" 0
.nrows = Get number of rows
for .row from 1 to .nrows
    .ext = Get value... .row .vel_ext_col
    if .ext = 1
        .r += 1
    elseif .ext = -1
        .f += 1
    endif
endfor

```

```

                endif
            endfor
            select .control
            Remove
        endproc

procedure append_points .table_id .time_col .value_col .grid .tier .tag$
    select .table_id
    .nrows = Get number of rows
    for .row to .nrows
        select .table_id
        .time = Get value... .row .time_col
        .value = Get value... .row .value_col
        select .grid
        Insert point... .tier .time '.tag$'newline$.value:2'
    endfor
endproc

procedure min .table_id .col
    select .table_id
    .nrows = Get number of rows
    .candidate = Get value... 1 .col
    .row = 1
    .index = Get row index... '.row'
    for .row from 2 to .nrows
        .entry = Get value... .row .col
        if .entry < .candidate
            .candidate = .entry
            .index = Get row index... .row
        endif
    endfor
    .min = .candidate
endproc

procedure max .table_id .col
    select .table_id
    .nrows = Get number of rows
    .candidate = Get value... 1 .col
    .row = 1
    .index = Get row index... '.row'
    for .row from 2 to .nrows
        .entry = Get value... .row .col
        if .entry > .candidate
            .candidate = .entry
            .index = Get row index... .row
        endif
    endfor
    .max = .candidate
endproc

procedure draw_results
#
# Raw and smoothed values are plotted on a log scale on the upper panel:
# the raw values in silver and the smoothed ones in grey. Values at tick
# marks are converted to the physical linear Hertz scale. The middle tick
# mark correspond to the median value of the smoothed F0 contour. The
# others correspond to the maximum and minimum values of the F0 contours
# (raw or smoothed, whatever has the broadest range).
#
# F0 velocity data values, expressed in Hz/dx, are plotted untransformed
# on the bottom panel. They indicate rate of F0 change in Hz over a fixed
# duration frame (referred to as dx) of 10 ms. The top and bottom
# horizontal tick marks correspond to the maximum and minimum values
# respectively. The middle horizontal dotted line marks the origin of the
# y-axis.
#
# In both panels green vertical lines correpond to maxima points and the
# red ones to minima points.
#
# The intervals defined by user-added boundaries at the IntervalTier
# specified at the initial GUI form are plotted in blue in the middle
# panel.
#
    ! Selecting some drawing parameters: y-axis limits
    ! F0 minimum and maximum values
    ! Obtain max and min values for raw contour

```



```

call min table 2
call max table 2
raw_min = min.min
raw_max = max.max
! Obtain max and min values for smoothed contour
call min table 3
call max table 3
sm_min = min.min
sm_max = max.max
! Compare them and select the broader range
if raw_min <= sm_min
    y_min = raw_min
else
    y_min = sm_min
endif
if raw_max >= sm_max
    y_max = raw_max
else
    y_max = sm_max
endif
! Get the smoothed F0 median value
select table
.temp_table = Extract column ranges... 3
.temp_table2 = To Table... row
y_median = Get quantile... F0-sm 0.5
select .temp_table
plus .temp_table2
Remove

! Draw the upper panel: raw and smoothed F0 contour overlaid
Erase all
Select outer viewport... 0 7 0 3
Axes... 0 file_end_time ln(y_min-5) ln(y_max+5)
Text top... yes Raw and smoothed %F_0
Text left... yes Hz
One mark left... ln(y_max) no yes no 'y_max:0'
One mark left... ln(y_min) no yes no 'y_min:0'
One mark left... ln(y_median) no yes no 'y_median:0'

select table
nrows = Get number of rows
for row to nrows
    time = Get value... row 1
    raw_F0 = Get value... row 2
    sm_F0 = Get value... row 3
    extremum = Get value... row 5
    ! Draw raw F0 value
    Paint circle (mm)... Silver time ln(raw_F0) 0.55
    ! Draw smoothed F0 value
    Paint circle (mm)... Grey time ln(sm_F0) 0.55
    if extremum = 1
        value = Get value... row 3
        Green
        Draw line... time ln(y_min) time ln(value)
    elseif extremum = -1
        value = Get value... row 3
        Red
        Draw line... time ln(value) time ln(y_max)
    endif
endfor

! Draw horizontal lines connecting S extremes
! to previous extreme

.temp = Extract rows where column... 5 "not equal to" 0
.nrows = Get number of rows
for .row to .nrows
    select .temp
    .ext = Get value... .row 5
    if .ext = 2 and .row > 1
        !
        if .row = 1
            !
            select table
            .begin = Get value... 1 1
            !
            select .temp
            !
            .end = Get value... .row 1
            !
            select table
            !
            .f0_begin = Get value... 1 3
            !
            select .temp

```

```

!                                     .f0_end = Get value... .row 3
!                                     else
!                                     .begin = Get value... .row-1 1
!                                     .end = Get value... .row 1
!                                     .f0_begin = Get value... .row-1 3
!                                     .f0_end = Get value... .row 3
!                                     endif
!                                     .mean = (.f0_begin + .f0_end)/2
!                                     Magenta
!                                     Draw line... .begin ln(.mean) .end ln(.mean)
endif
endfor
select .temp
Remove

! Draw the middle section: user-provided intervals
original = Read from file... 'grid$'name$.TextGrid
interval_tier = Extract tier... 'segmentation'
draw_grid = Into TextGrid
Select outer viewport... 0 7 1 3.5
select draw_grid
Blue
Draw... 0 0 no yes no
select original
plus interval_tier
plus draw_grid
Remove

! Bottom panel: F0 velocity contour and extreme points
! Valores máximos e mínimos para F0
call min table 4
call max table 4
y_min = min.min
y_max = max.max
Select outer viewport... 0 7 3 6
Axes... 0 file_end_time y_min-0.1 y_max+0.1
Text bottom... yes 'subtitle$'
Text left... yes Hz/dx, dx=10 ms
One mark left... 0 yes yes yes
One mark left... 'y_max:2' yes yes no
One mark left... 'y_min:2' yes yes no
nrows = Get number of rows
for row to nrows
time = Get value... row 1
extremum = Get value... row 6
point = Get value... row 4
Paint circle (mm)... Grey time point 0.55
if extremum = 1
value = Get value... row 4
Green
Draw line... time y_min time value
elseif extremum = -1
value = Get value... row 4
Red
Draw line... time value time y_max
endif
endif
endfor
Select outer viewport... 0 7 0 6
Draw inner box
Select outer viewport... 0 0.5 6 6.5
Black
endproc

procedure write_report

! F0 smoothing factor
.f0_sm = factor

! Was F0 velocity contour smoothed?
if smooth_derivative = 1
.f0_vel$ = "T"
else
.f0_vel$ = "F"
endif

! F0 velocity smoothing factor
if .f0_vel$ = "T"
.f0_vel_sm$ = "'factor2'"

```

```

else
    .f0_vel_sm$ = "NA"
endif

! Threshold method chosen
.method = threshold
if .method = 1
    .method$ = "Glissando threshold"
elseif .method = 2
    .method$ = "+1 semitone"
else
    .method$ = "None"
endif

! Glissando threshold value
if .method = 1
    .thr$ = "'gtr'"
else
    .thr$ = "NA"
endif

! Correlation coefficient between smoothed and raw F0 values
select table
.temp_table = Extract column ranges... 2 3
.corr_table = To Correlation
.r_sqrd = Get value... 1 2
.r_sqrd = .r_sqrd^2
select .temp_table
plus .corr_table
Remove

! Median and maximum error of smoothed F0 contour
select table
.temp_table = Extract column ranges... 2:3
select .temp_table
Insert column (index)... 3
Insert column (index)... 4
Set column label (index)... 3 error-Hz
Set column label (index)... 4 error-st
Formula... if col = 3 then abs(self[1] - self [2]) else self fi
Formula... if col = 4 then abs(hertzToSemitones(self[1]) - hertzToSemitones(self[2])) else self
fi

.temp_table2 = To Table... row
.median_Hz = Get quantile... error-Hz 0.5
.median_st = Get quantile... error-st 0.5
.max_Hz = Get maximum... error-Hz
.max_st = Get maximum... error-st
select .temp_table
plus .temp_table2
Remove

! Collapse S chains option
if collapse_chains = 1
    .collapse$ = "T"
elseif collapse_chains = 0
    .collapse$ = "F"
elseif threshold = 3
    .collapse$ = "NA"
endif

! Print report to Praat Info window
printline Parameters Report'newline$'-----
printline file:'newline$'      'name$'
printline F0 smoothing factor:'newline$'      '.f0_sm' Hz
printline F0 velocity smoothing:'newline$'      '.f0_vel_sm$' Hz
printline raw-smoothed F0 R^2:'newline$'      '.r_sqrd:2'
printline median smoothing error:'newline$'      '.median_Hz:2' Hz, '.median_st:2' st
printline maximum smoothing error:'newline$'      '.max_Hz:2' Hz, '.max_st:2' st
if .method = 1
    printline threshold method:'newline$'      '.method$', '.thr$' ST/s
else
    printline threshold method:'newline$'      '.method$'
endif
printline collapse S chains:'newline$'      '.collapse$'
printline script execution time:'newline$'      'clock:2' s

! Write report file
if script_report = 1

```

```

        report$ = grid$ + name$ + "_log.txt"
        filedelete 'report$'
        line$ =
"name$"tab$.f0_sm"tab$.r_sqrd:2"tab$.median_Hz:2"tab$.median_st:2"tab$.max_Hz:2"tab$.m
ax_st:2"tab$.f0_vel"tab$.f0_vel_sm"tab$.method"tab$.thr"tab$.collapse$"
        line$ >> 'report$'
    endif
endproc

```

Anexo IV


```

"ooTextFile"
"ExperimentMFC 5"
stimuliAreSounds? <yes>
stimulusFileNameHead = "tone_test/"
stimulusFileNameTail = ".wav"
stimulusCarrierBefore = ""
stimulusCarrierAfter = "tone"
stimulusInitialSilenceDuration = 0.5 seconds
stimulusMedialSilenceDuration = 1.2 seconds
numberOfDifferentStimuli = 17
    "agrega_test" "isso nn, isso não agrega NADA, né?"
    "antidopping_test" "ô André, o Brasil anda MUITO atrasado nessa questão do anti-dopping
hein!"
    "comendo_test" "então o problema não está do que você está COMENDO ou deixando de comer"
    "gol_test" "minha vida é muito mais feliz graças àquele GOL"
    "goleiro_test" "o GOLEIRO ele é cargo de confiança do treinador"
    "harrypotter_test" "tá namorando o cara que é dono do castelo de HARRY POTTER cara"
    "imagem_test" "veja bem, nós temos uma IMAGEM, subretudo na moda, né?"
    "jornal_test" "parece mentira, mas eu não li JORNAL hoje, não deu tempo"
    "mariadapenha_test" "os CASOS da lei Maria da Penha são relacionados à violência
doméstica"
    "medo_test" "né, esse medo, como eu te disse, vira MANIA, essa mania vira toc"
    "antidopping_test2" "ô André, o Brasil anda muito ATRASADO nessa questão do anti-dopping
hein!"
    "comendo_test2" "então o problema não está do que você está comendo ou deixando de COMER"
    "imagem_test2" "veja bem, nós temos uma imagem, subretudo na MODA, né?"
    "goleiro_test2" "o goleiro ele é cargo de confiança do TREINADOR"
    "jornal_test2" "parece mentira, mas eu não li jornal hoje, não deu TEMPO"
    "mariadapenha_test2" "os casos da lei Maria da Penha são relacionados à VIOLÊNCIA
DOMÉSTICA"
    "medo_test2" "né, esse medo, como eu te disse, vira mania, essa mania vira TOC"
numberOfReplicationsPerStimulus = 2
breakAfterEvery = 0
randomize = <PermuteBalancedNoDoublets>
startText = "Esse é um experimento de percepção.
Depois de ouvir as gravações
classifique o tom palavra destacada de acordo com sua percepção.

Clique na tela para começar."
runText = ""
pauseText = ""
endText = "O experimento acabou. Obrigada!"
maximumNumberOfReplays = 3
replayButton = 0.80 0.90 0.10 0.20 "Repetir" ""
okButton = 0 0 0 0 "" ""
oopsButton = 0 0 0 0 "" ""
responsesAreSounds? <no> "" "" "" "" 0 0
numberOfDifferentResponses = 3
    0.20 0.30 0.10 0.20 "ALTO" 10 "" "1"
    0.35 0.45 0.10 0.20 "NÃO SEI" 10 "" "0"
    0.50 0.60 0.10 0.20 "BAIXO" 10 "" "-1"
numberOfGoodnessCategories = 0

```


Anexo V


```

# WritingDownTextGrid.psc
# Script implemented by Plinio A. Barbosa (IEL/Unicamp) in the framework
# of the Swed-Bra collaboration.
# Date: Dec., 2010
form Input file
    word Tg_suffix *.TextGrid
    integer Tiernumber 1
    word Outputfile textout.txt
endform
filedelete 'outputfile$'
Create Strings as file list... list .\'tg_suffix$'
numberOfFiles = Get number of strings
for i to numberOfFiles
    select Strings list
    tg$ = Get string... 'i'
    Read from file... 'tg$'
    tgname$ = selected$("TextGrid")
    nint = Get number of intervals... 'tiernumber'
    fileappend 'outputfile$' 'tgname$' 'newline$'
    for j to nint
        temp$ = Get label of interval... 'tiernumber' 'j'
        if temp$ <> ""
            fileappend 'outputfile$' 'temp$'
        endif
    endfor
    fileappend 'outputfile$' 'newline$'
endfor

```

